Journal of
Experimental Psychology

VOL. VII., No. 1 February, 1924

SPEED, ACCURACY AND CONSTANCY OF RESPONSE TO VISUAL STIMULI AS RELATED TO THE DISTRIBUTION OF BRIGHTNESSES OVER THE VISUAL FIELD [1]

BY H. M. JOHNSON

Ohio State University

This paper reports the results of 14,000 measurements of the times required for selective reaction to two visual stimuli by five trained subjects. The brightness of the test-field and the intensity of the stimuli were kept practically constant throughout the experiment. The variable factor tested was the relative brightness of the field surrounding the test-object.

The purpose of the study was to secure additional evidence as to the influence of the variable factor on visual performance, for use by those responsible for the determination of a policy with respect to the lighting of a large plant devoted to the manufacture of rubber goods.

The considerations which led to the choice of this variable rather than of certain others may be understood by reference to a summary of the work of other investigators on other aspects of illumination as related to vision.

Due principally to the advertising propaganda of the manufacturers of window-sash and electric lamps, and to the influence of this propaganda on the engineering mind, the two aspects of illumination which have almost monopolized

[1] The experimental work reported in this paper is from the research laboratory, Engineering Department, The B. F. Goodrich Company. It was prepared for publication in the psychological laboratory, University of Minnesota.

public attention are (1) the absolute intensity of illumination and (2) the wave-length composition of the light.

The Dependence of Vision on the Absolute Brightness of the Visible Surface

The absolute brightness (B),[2] of a visible surface is determined by the product of the absolute density of illumination (E) which falls upon it, and the coefficient of reflection (R') in the direction of the eye. In general $B = ER'/\pi$. If R' be constant $B \propto E$. Engineers have tended to reduce their lighting questions to the form, "what illumination does one need in order to see?" This formulation gives slight consideration to the factor of reflectivity and little or no consideration to that of adaptation. The recent invention of the 'foot-candle meter,' which utilizes the principle of the grease-spot photometer in a portable form, capable of manipulation by a novice, has had the unfortunate effect of strengthening this tendency.

Carried to extremes, as in the work of Durgin (10), such a viewpoint leads to disregard, not only of the operation of other factors than absolute illumination in vision, but also of the operation of other factors than vision in productive labor. In Durgin's report we have a comparison of war-time production before and after an increase in the general level of illumination, with all the improvement referred to that increase, and without entertainment of such minor questions as the rate of supply of raw material, the rate of pay, the basis of pay, the health of the employees, the seasonal changes in productivity, the extraordinary effects of patriotic appeal, the 'work-or-fight' order, and the like. In this discussion, we shall bear in mind that the chief use of illumination—to the worker—is to enable him to see, and shall confine our review to studies made on the influence of absolute brightness on vision. Of the visual functions which have been referred to this variable, the first which we shall consider is that of the perceptibility of differences in brightness.

[2] If by E we designate the illumination in foot-candles, by R' the reflectivity the direction of the eye, and by B' the brightness expressed in millilamberts, then $B' = 1.077ER'$.

Brightness—Difference Sensitivity and Absolute Brightness

In 1888 Koenig and Brodhun (19), in 1913 Cobb and Geissler (2), and in 1916 Cobb (4) published very definite results on this point. Koenig's data are most readily interpreted in the form to which they were reduced by Nutting (21). They show that the relative differential threshold $\delta B/B$ for absolute brightness is practically constant between 1 and $(10)^3$ millilamberts for all wave-lengths. It increases rapidly as the absolute brightness is reduced below one millilambert, the increase being more rapid for the longer wave-lengths than for the shorter because of Purkinje's effect. Cobb and Geissler's (2) and Cobb's (4), results on tungsten light are in full agreement with Koenig and Brodhun's (19) under adaptation to the brightness of the field. Under that condition no appreciable gain in this form of sensitivity results from increasing the brightness of the viewed surface much above 1 millilambert.

'Visual Acuity' and Absolute Brightness

'Visual acuity,' *i.e.*, the reciprocal of the visual angle expressed in minutes which yields liminal separations of two portions of the visual field, has been related to the absolute mean brightness of the test-field by Koenig (20) Cobb and Geissler (2), Cobb (3), Dunlap (8), and Ferree and Rand (14), (16). The several sets of measurements were made under different conditions, which accounts in part for the diversity of the results. Koenig and Ferree and Rand used figures impressed on gray paper in printer's ink; the former utilizing Snellen's 'hooks,' and the latter Landolt's broken rings. Their test-objects were viewed in surroundings whose brightnesses were of approximately the same order as that of the field. Cobb and Geissler and Cobb used a diffusing-glass viewed through a pair of crossed gratings in an Ives-Cobb mounting, with an exposure-time of 3 seconds. The results herein presented were obtained in dark surroundings. If the surrounding field was given a finite brightness not exceeding that of the test-object, fixation was facilitated, and absolute

'acuity' enhanced; but the general course of the curve was not greatly changed. I chose the measurements made in dark surroundings because they yielded a larger number of points for plotting. Dunlap used a uniformly illuminated slit viewed through a tube containing a rotatable crystal of Iceland spar, which gave two images of the slit of variable displacement. The entire stimulus-field was of foveal extent. As he employed the method of adjustment the time of exposure was not controlled.

Fig. 1 presents a graphic summary of the results of the

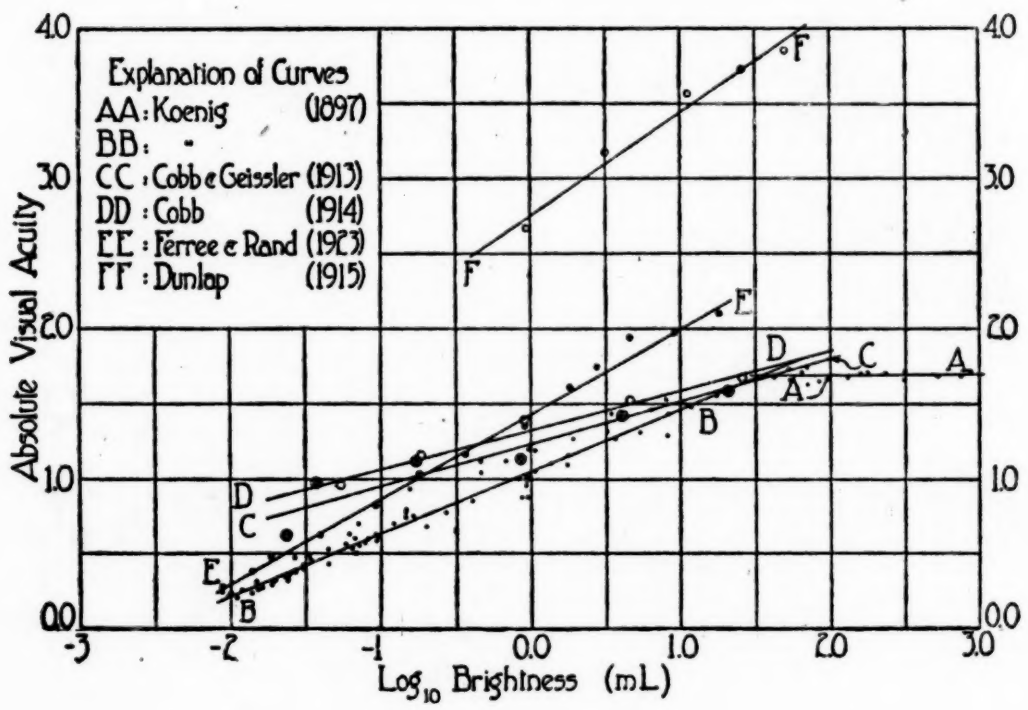

Fig. 1

several authors. From inspection it appears that within certain limits of absolute brightness the following relationship holds with fair accuracy:

$$A = c + k \log B, \tag{1}$$

A representing absolute acuity, B absolute brightness in millilamberts, and c and k constants obviously $A = c$ when $B = 1$. Designating this value of A by A_0, we may express relative acuity A/A_0 by

$$A/A_0 = (k/c) \log B \tag{2}$$

and thereby reduce all the measurements to a common basis for comparison. Table I gives the constants for the above

equations and the limits within which they hold for each set of measurements.

TABLE I

CONSTANTS AND LIMITS FOR RELATIONSHIP BETWEEN VISUAL ACUITY AND BRIGHTNESS

A = absolute visual acuity = 1 ÷ liminal visual angle in minutes.
A_0 = value of A corresponding to one millilambert.
B = brightness in millilamberts.

$$A = c + k \cdot \text{Log}_{10} B$$
$$A/A_0 = 1 + (k/c)\ \text{Log}_{10} B$$

Author	Limits (mL)	c	k	k/c
Koenig	0.01 ... 43.5	1.05	0.415	0.394
Koenig	40. ... 1,000.	1.69	0.000	—
Cobb & Geissler	0.02 ... 21.	1.23	0.282	0.229
Cobb	0.06 ... 26.	1.33	0.262	0.197
Dunlap	0.9 ... 51.	2.76	0.687	0.249
Ferree & Rand	0.1 ... 18.	1.44	0.573	0.397

Two facts are noteworthy. First, as Koenig and subsequent writers have observed, the relationship between acuity and log B is not continuous for high and low brightnesses and the scatter in the lower region is high. The relationship clearly begins to fail at brightnesses greater than 40 millilamberts, increases above 180 millilamberts resulting in actual impairment.

Secondly: as may be noted from inspection of the different values of k/c (the constant which determines the slope of the curve of relative acuity A/A_0 plotted against log B), there is considerable diversity. Ferree and Rand's effect, as indicated by k/c is 1.01 times Koenig's, 1.59 times Dunlap's, 1.73 times Cobb and Geissler's and 2.02 times Cobb's. The last three ratios seem rather large. I would be inclined to attribute them to the short duration of exposure used by Ferree and Rand, which one would expect to sensitize the test, except for my recollection of my own experiences as observer, and my conversations with the other observers, in Cobb's and Dunlap's studies. While Cobb used an exposure of 3 seconds, we tended to use only the first small fraction of the time. If the eye were in satisfactory fixation at the instant of exposure the pattern was instantly visible though it might

the separate factors in visual performance, they need to be interpreted with care. The comment on Ferree and Rand's findings may be applied *a fortiori* to those of the other authors.

Ferree and Rand's study brings out one other interesting point which requires attention here. One of their subjects (whose refraction, presumably, was incapable of improvement by stock lenses) was tested without glasses and also with cylindrical lenses of different strengths, to render the images slightly astigmatic. The strongest lens used, namely, + 0.25 D., gave the greatest effect. The 'visual acuity' of the eye when astigmatized to this degree was much lower relatively to that of the normally used eye at low brightnesses than at moderate and high brightnesses. This disparity diminishes rapidly as a brightness of 0.46 millilambert is approached. Between 0.46 and 33 millilamberts, the ratio between the acuity A_a of the astigmatized eye and that of the normally used eye (A_n) is quite accurately expressed by the formula

$$A_a/A_n = 0.946 + 0.022 \operatorname{Log}_{10} B \qquad (3)$$

From this relationship it is clear that within the range of brightnesses under which visual work is ordinarily done, the effect of absolute brightness on the 'acuity' of an eye which has been astigmatized to 0.25 D. is but very little greater than on the 'acuity' of the normal eye. A contrary belief now seems to be prevalent among lighting experts. Certain of them, in consultation with an architect employed by a school board, recently cited the work of these authors in support of their contention. Apparently, their impressions came from a rather casual inspection of the authors' Fig. 2, without a careful examination of the accompanying text.

Liminal Duration of Exposure and Absolute Brightness

In 1919, following an extended analysis and discussion of existing tests, Cobb suggested that since the eye is ordinarily at rest for very brief periods (0″.07 to 0″.1 at the least, according to Dodge, and 0″.11 to 0″.19 in the average, according to Huey), and since in such work as reading the

most important parts of the field are usually imaged on the parafoveal retina rather than on the fovea itself, the ordinary type of measure of visual performance was inappropriate. The essential characteristics of an appropriate test were to be: a stimulus (1) of small, but supraliminal area, presenting (2) a constant difference in brightness with respect to its background, (3) imaged slightly off the fovea, and (4) of variable duration of exposure. The threshold of effectiveness was to be measured in terms of this duration. These considerations were published in 1920 (5). In the year 1919–20 Cobb and Loring (6) made a preliminary investigation of the merits of this method as an exhibitor of individual differences and standardized it. Cobb then applied it to an investigation of a problem bearing on visual demands of aviation (7). He found important differences among the individual observers, which were correlated to a practically insignificant degree with the results of conventional ophthalmologic measurements. From this it appears that the test involves a different complex of variable visual factors from those covered by the conventional tests.

Later (8) Cobb applied this test to the effect of the absolute brightnesses of the test-field. Seven observers were employed between limits of 7.8 and 107 millilamberts, and three observers between limits of 1.05 and 107 millilamberts; the measurements within the lower ranges being made after those within the higher. The reciprocals of the liminal durations of exposure (called by the author the 'speeds of retinal impression') are resumed graphically in curves *AA* and *BB* of Fig. 3. The points plotted are the weighted geometric means of the two groups of observers. Within the given brightness limits these speeds (S) are fairly expressed by the equation

$$S = c + k \cdot \log B \qquad (4)$$

c and k being constants.

Obviously, since $S_0 = c$ when $B = 1$, the relative speed

$$S/S_0 = 1 + (k/c) \log B \qquad (5)$$

a result which is shown graphically in curves *AA* and *BB* Fig. 4.

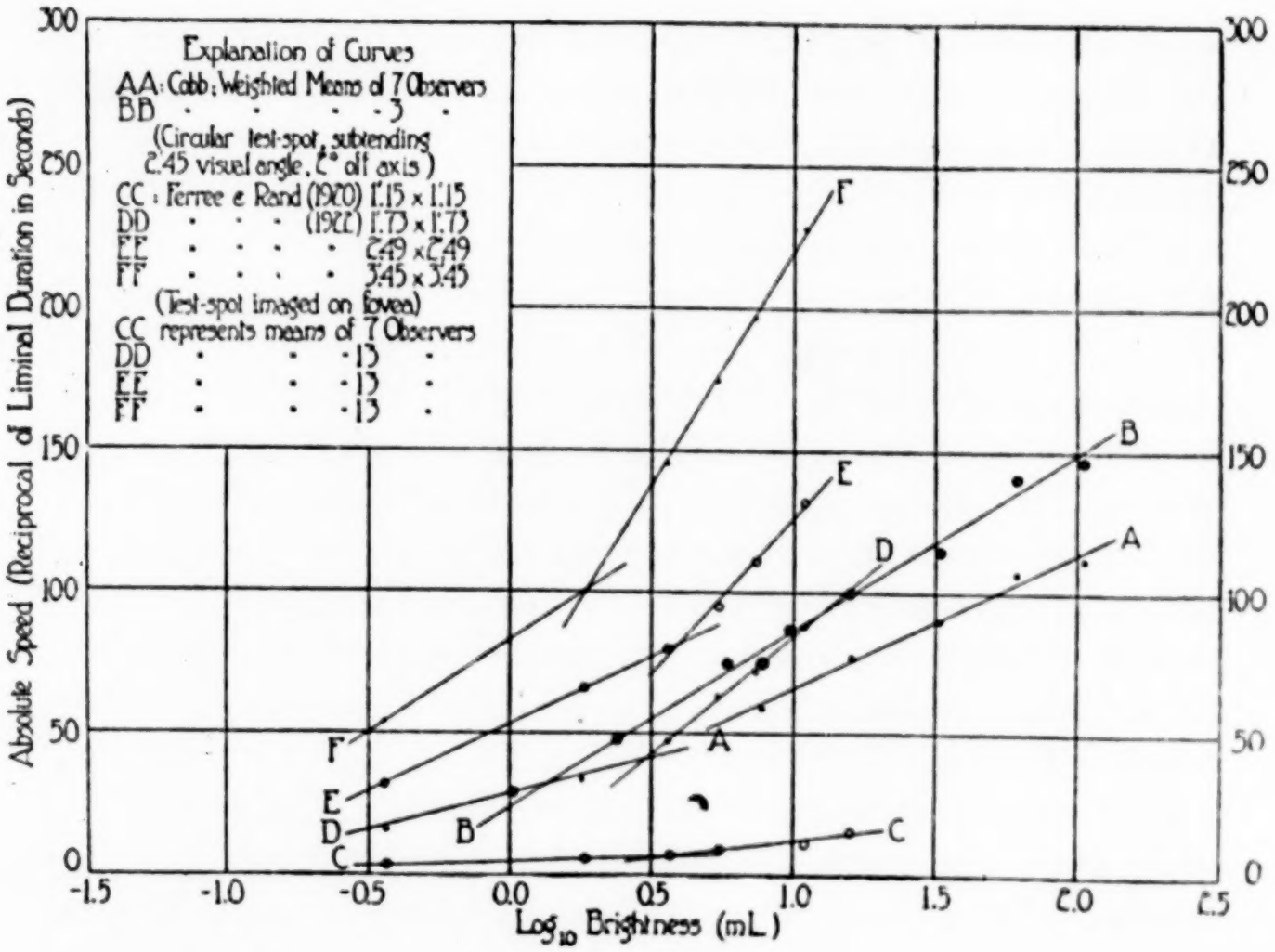
Explanation of Curves
AA: Cobb; Weighted Means of 7 Observers
BB · · · 3 ·
(Circular test-spot, subtending 2'.45 visual angle, 2° off axis)
CC: Ferree & Rand (1920) 1'.15 × 1'.15
DD · · · (1922) 1'.73 × 1'.73
EE · · · · 2'.49 × 2'.49
FF · · · · 3'.45 × 3'.45
(Test-spot imaged on fovea)
CC represents means of 7 Observers
DD · · 13 ·
EE · · 13 ·
FF · · 13 ·
Absolute Speed (Reciprocal of Liminal Duration in Seconds)
Log10 Brightness (mL)
300
250
200
150
100
50
0
-1.5
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
2.5
F
E
D
B
A
C

FIG. 3

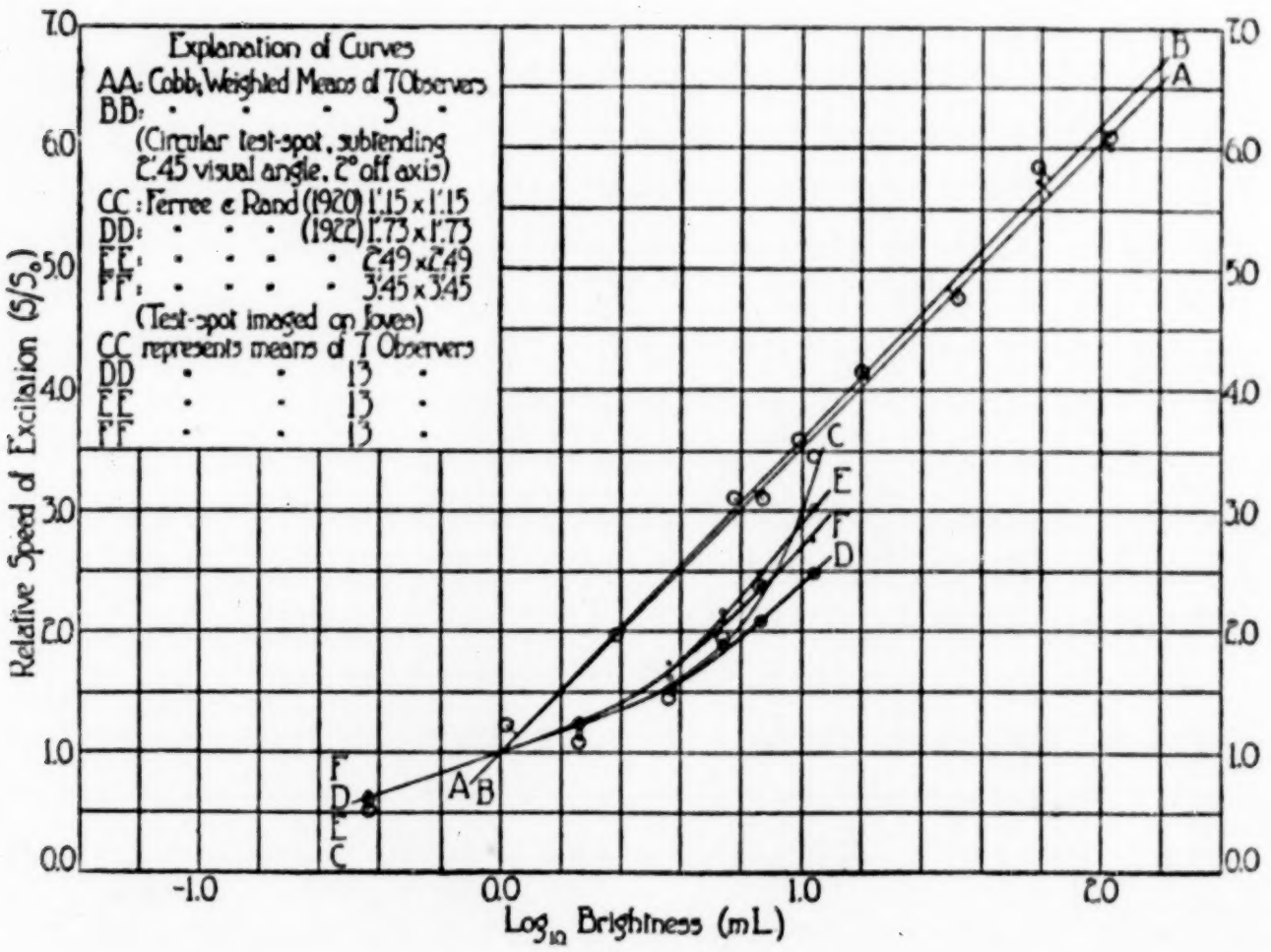
Explanation of Curves
AA: Cobb; Weighted Means of 7 Observers
BB: · · · 3 ·
(Circular test-spot, subtending 2'.45 visual angle, 2° off axis)
CC: Ferree & Rand (1920) 1'.15 × 1'.15
DD: · · · (1922) 1'.73 × 1'.73
EE: · · · · 2'.49 × 2'.49
FF: · · · · 3'.45 × 3'.45
(Test-spot imaged on fovea)
CC represents means of 7 Observers
DD · · 13 ·
EE · · 13 ·
FF · · 13 ·
Relative Speed of Excitation (S/S0)
Log10 Brightness (mL)
7.0
6.0
5.0
4.0
3.0
2.0
1.0
0.0
-1.0
0.0
1.0
2.0
B
A
C
E
F
D

FIG. 4

From the author's discussion one is given to understand that in his opinion the upper and lower limits of the range of brightnesses which he employed may be near the limits of applicability of the law.

The values of the several constants are shown in the first two rows of Table 2.

Meanwhile, Ferree and Rand (14), (15) published the results of an investigation of the effects of absolute brightness as indicated by a test which has some features in common with Cobb's and some features which are distinctive. Their test object was a Landolt broken ring, imaged on the fovea. The angular extent of the break in the ring was varied between $1'.15 \times 1'.15$ and $3'.45 \times 3'.45$ of visual arc. The results are resumed in curves *CC*, *DD*, *EE* and *FF* of Fig. 3, in which the reciprocals of the liminal durations are plotted against the logarithm of the brightness. Each point represents the unweighted arithmetic means of 7 or 13 observers, as the case may be. These reciprocals the authors call the 'speed of discrimination'—an expression which I regard as unfortunate.

While the points in each case could be fitted to a single smooth curve it seems more appropriate to regard the relationship between 'speed' and brightness as rectilinear and discontinuous for higher and lower brightnesses. That is to say, between definite brightness limits

$$S' = c' + k' \log B, \tag{6}$$

and

$$S'/S_0' = 1 + (k'/c') \log B \tag{7}$$

The constants for the two limbs of the several curves derived by least squares, are shown in the lower part of Table II.

Fig. 4, curves *CC*, *DD*, *EE* and *FF* show the relationship between relative speed S_0' and brightness, S_0' being the speed corresponding to a brightness of one millilambert as derived from the equation of the lower limbs of the curves. In this figure I have simply plotted the points and have drawn in the second degree curve which could be most easily fitted to

them by a lowly order of draftsmanship. These curves are intended simply as aids in locating the plotted points.

TABLE II

CONSTANTS FOR RELATIONSHIPS BETWEEN "SPEED" (S) AND BRIGHTNESS (B)

$$S = c + k \log_{10} B$$
$$S/S_0 = 1 + (k/c) \log_{10} B$$

Author	Dimensions of Stimulus	c	k	k/c	Range of Applicability (Millilamberts)
Cobb	circle, D =				
(7 observers)	2′ 45	18.7	47.4	2.53	7.8 ..107
(3 observers)	2′ 45	24.1	62.4	2.59	1.05..107
Ferree & Rand	1′ 15×1′ 15	4.36	4.04	0.93	0.37... 3.67
Ferree & Rand	1′ 15×1′ 15	− 2.48	15.55	*	3.67...10.9
Ferree & Rand	1′ 73×1′ 73	28.99	25.88	0.89	0.37... 3.67
Ferree & Rand	1′ 73×1′ 73	0.09	84.06	*	3.67...10.9
Ferree & Rand	2′ 49×2′ 49	53.12	48.05	0.905	0.37... 3.67
Ferree & Rand	2′ 49×2′ 49	16.04	110.3	*	3.67...10.9
Ferree & Rand	3′ 45×3′ 45	82.74	65.62	0.79	0.37... 1.83
Ferree & Rand	3′ 45×3′ 45	55.38	164.00	*	1.83...10.9

S = reciprocal of liminal duration of exposure in seconds
S_0 = S corresponding to 1 mL

Except for the case in which the stimulus was 1′.15 × 1′.15 in extent, and therefore near a liminal value, the upper portions of these curves are almost parallel to Cobb's, though lower. Had a different reference brightness than 1 millilambert been chosen for S_0′ the correspondence would have appeared more striking, but the desirability of adhering to a uniform procedure outweighed this consideration.

The results of Cobb and of Ferree and Rand agree in showing that what they call 'speeds'—of whatever kind—vary to an important degree and in a highly consistent manner, with brightness. But no single form of presentation exhibits all the important characteristics of their data, and with some of them omitted from exhibition a layman may be erroneously impressed by the rest.

For example: an engineer inspecting curve *BB*, Fig. 3 or 4, might reason that since the reciprocal of the time corre-

* This value not applicable.

sponding to 10 millilamberts is 3.6 times as great as that corresponding to 1 millilambert, therefore the subject 'sees 3.6 times as quickly' in the former case as in the latter. His predilection for that interpretation is strengthened by the use of the term 'speed' in any connection, but it is especially reinforced by the designation 'speed of discrimination' employed by Ferree and Rand.[1] (I haven't yet seen such an interpretation of these results used in rationalizing the 'booster' propaganda for 'filling sockets' or 'filling sashes,' but I expect it to be so used within the near future.)

For a proper perspective of the results the reader should recall that the thing originally measured was time: that is to say, the duration of exposure necessary for a common degree of effectiveness. This value does *not* represent the time required for discrimination. The term 'discrimination' is used by modern psychologists to designate the making of a *differential* or *selective response*—*i.e.*, usually a muscular contraction—which is conditioned by one stimulus and not by another. The average time required for the simplest sort of selective or discriminatory response to a visual stimulus is under the most favorable conditions somewhere near 0."27. The term is also legitimately used to designate the process of being 'conscious of' or 'aware of' an object as distinguished from another object. There is not the slightest evidence that the process of 'becoming aware of' the object which is 'discriminated' precedes the differential response of the muscles; indeed the contrary is implied in the arguments of certain psychologists of repute who are still interested in the study of consciousness.

Now the liminal times which these authors determined are the durations of exposure of the stimulus which are necessary merely to *initiate* a selective response; and they constitute a very small part of the total time required for 'discrimination.'

[1] I hasten to add that an advance notice of Dr. Cobb's results, printed for private circulation among the technicians and salesmen of the National Lamp Works, contained a report of an interview with him in which he warned against such misinterpretation and gave his reasons in detail. The views of Dr. Ferree may be gleaned from the printed discussion appended to (14) and (15).

Practically the question should be formulated thus: how much time is saved by a given increase of absolute brightness, and what use may be made of the salvage? The first part of this question may be answered by reference to Fig. 5.

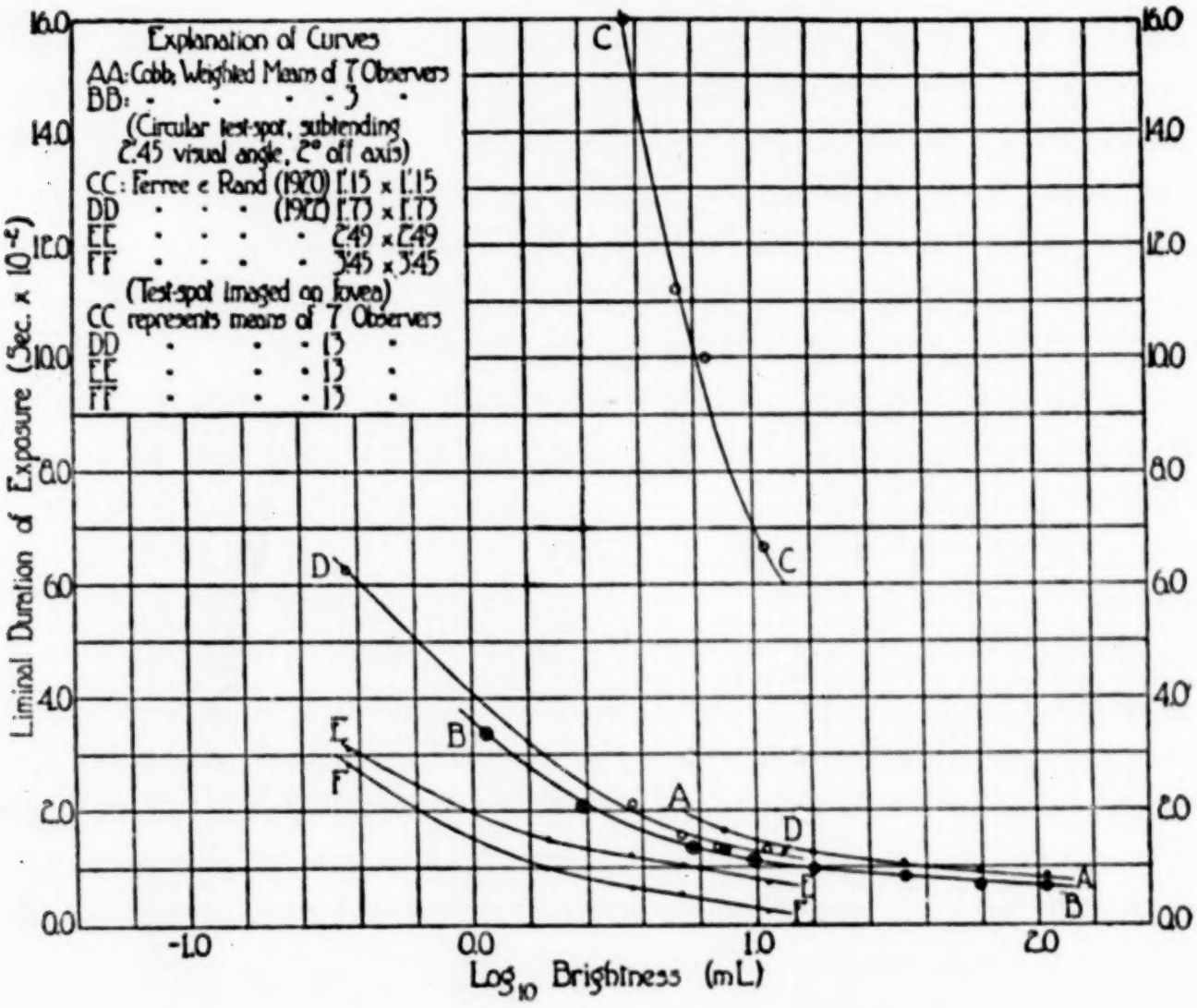

FIG. 5

In this figure, Curve *BB*, for example, shows the liminal time for 1 *mL* to be about 0."036; for 10 *mL*, 0."011; for 100 *mL*, 0."007. The first difference merits attention; the absolute size of the latter is pretty small. Doubling an original brightness of 1 millilambert gives a saving of 0."013, a value which is important or not depending on the value with which it is to be compared.

With respect to the total time ordinarily required for selective response to one of two stimuli it is about five per cent.; with respect to the time required for selective response to one of eight stimuli (as used by Ferree and Rand), it is about two per cent. In neither case would it be negligible if

one knew that the total time would be consistently affected to that degree. As compared with the normal resting period of the eye (say, 0″.1) it is 13 per cent. If the normal rate of eye movement had to be interfered with by so large a prolongation of the resting period, the effect might be serious; but if the duration of exposure, as in this instance, is already but 36 per cent. of the resting period the advantage of shortening the necessary duration is not immediately obvious. More facts need to be known for satisfactory interpretation.[1]

There is one possibility which ought to be considered, and which the authors should be able to settle without further experimentation. The minimum duration of exposure whose effectiveness is practically perfect might be more enlightening than the liminal duration. If this time for a given brightness

[1] As the technical reader will recall, the serial eye-movements are very complex and delicately coördinated responses. This type of movement is called saccadic, possibly from its resemblance to that of a horse whose reins are jerked tightly after each step. At any rate it consists of a series of jerky movements, separated by brief pauses. During the jerky movement the subject is blind—presumably because the neural impulses aroused in the retina are somehow blocked. Furthermore these series of movements are at least partially 'circular':—that is, the contraction of one group of muscles tends to excite the contraction of the next group in the temporal order, without much aid from other stimuli. As in the serial reflexes in walking, the responses tend to be made in a certain rhythm; and interference with this rhythm makes the coördination more difficult. For example, one finds it far more wearisome to walk on the ties of a railroad track than on the rail, because in the latter case the length and cadence of step to which one is habituated may be used; while in the former each step must be adjusted to the separation of the ties, and hence is largely a response to other stimuli than those of muscular contraction in the legs and trunk themselves, and is interfered with by the tendency to respond to these contractions. Similarly as to eye-movement; if the habitual rhythm is interfered with, the external muscles may work in mutual antagonism and hence suffer far greater fatigue. The reflex of winking may also be inhibited. It is possible that most of the discomforts vaguely referred to as 'eyestrain' proceed from these sources.

If the visual task requires a longer duration of exposure than the habitual fixational pause, it is easy to see how the trouble arises, for one group of the external muscles must struggle to hold the eye in too long a fixation against the action of an antagonistic group which have been stimulated by the previous activity to move the eye along. If a very short duration of exposure is adequate for retinal excitation, the habitual resting period may not be shortened without interference; if the fixation is prolonged to its normal length the eye will be overstimulated and abnormally intense after-images will be aroused, unless the mechanism is somehow rendered less sensitive. I would not contend that the habitual rhythm of eye-movement could not be accelerated by the formation of a new system of habits, but obviously such modification is subject to limitations as well as difficulties.

is a major fraction of the resting period for a stimulus of supraliminal area, it would seem advisable to increase the brightness materially. Just what factor of safety should be adopted cannot be determined out of hand.

The curve I chose for discussion happens to be Cobb's but the discussion can be transferred, in some cases with increased emphasis, to all the others but one. That one (*CC*) represents Ferree and Rand's smallest stimulus which, except at the highest brightness required an exposure as long as, or in excess of, the normal resting period.

In this connection should be mentioned another interesting test of Ferree and Rand's, made on the liminal duration of exposure for designation of the position of the breaks in three Landolt rings, successively exposed in the line of vision, and at distances of 20, 600 and 20 cm. respectively. The liminal durations varied between 1″.36 and 2″.44 between brightnesses of 0.37 and 10.9 millilamberts. The reciprocals of these times are called the 'speeds of adjustment' (S'''), and are a rectilinear function of the logarithm of the brightness. From the equation of the 'best fitting' line

$$S''' = 0.515 + 0.227 \log_{10} B, \tag{8}$$

$$S'''/S_0''' = 1 + 0.441 \log_{10} B. \tag{9}$$

The authors refer to this test as a relatively insensitive index of the effect of brightness but from equation (9) it appears that it is about 10 per cent. more sensitive than their test of 'acuity.' The break in the more distant ring subtended a visual angle of 1′.4 × 1′.4, which is rather near the threshold. This fact may account for a considerable part of the large value of k/c.

Reaction-time and Absolute Brightness

I might cite a number of references to fairly ancient work done on the relation between mean brightness of the test field and the time required for simple and selective reactions; but their perusal would be unfruitful, for two reasons: first, because the absolute brightnesses were not measured nor were data presented which permit of their

computation; and secondly, because the results are invariably negative or inconsistent. Elsewhere (18) I have discussed some of the reasons for their divergence.

More recently Elliott (13) measured the times required for what she calls 'cognitive reactions' to visual stimuli. These were numbers of four digits each printed on paper whose reflectivity was not given. The test object was illuminated by tungsten, solar and mercury-vapor light of varying densities. The results taken at face value indicate great superiority of mercury light over solar and of solar over tungsten; and show large diminutions in time with increases in illumination up to 10 foot-candles. They are now being vigorously exploited by partisans of the mercury-vapor lamp (11), (12). I have elsewhere shown (18) that the 'cognitions' were expressed by a different muscular response from the one which stopped the chronoscope, and that there is no ground for assuming that the two acts coincided or that there was a constant temporal interval between them. The results therefore cannot be profitably considered further.

The above discussion summarizes the most important results obtained by objective and reliable methods on the relation between visual performance and absolute brightness. Some of the data were not extant in 1920, when the present work was undertaken.

Wave-length Composition

The problem of the effect of wave-length composition of light on visual performance is of great interest and now demands a careful attack.

It was excluded from consideration however because in its practical form it was reduced to a question of the relative merits of mercury-vapor and tungsten sources. While it was recognized that this question is an open one with respect to certain types of operation, yet in this particular institution, the decision had been made in favor of tungsten and a change was considered inadvisable on economic grounds.

On the basis of information then available we assumed that a mean field-brightness of one or two millilamberts was

adequate for all work well above threshold dimensions, provided adaptation to that brightness could be maintained. This opinion I still hold, subject to modification in the face of adequate evidence.

Distribution of Brightnesses and Visibility

If a surface is illuminated to an intensity E, the brightnesses of its several elements will be proportional to their several coefficients of reflection in the direction of the eye. If only E is varied, the only effect on vision will be to alter these several brightnesses in the same proportion, as long as the change does not exceed the limits of brightness to which a constant sensory adaptation is possible. From the work of Koenig on the perceptibility of relative differences in brightness, it appears that these limits are in the neighborhoods of one and 1,000 millilamberts, respectively.

Within these limits the sensitivity to relative differences in brightness is practically constant, so that under stable adaptation it matters little whether the absolute level of brightness be high or low as long as these limits are not exceeded. This applies to stimuli well above liminal sizes. If their areas be excessively small, visual acuity becomes a factor. The effect of absolute brightness on visual acuity may be due in large part to the tendency of the pupil to assume a smaller diameter under the higher illuminations, and hence to diminish the effect of spherical aberration in the refracting media and thereby to improve the definition of the image. At the same time the diminished size of the pupil tends to reduce the resolving power of the refractory system and thereby to impair definition, but it would seem better to expect the contrary effect to prevail among the majority of the population, because most eyes suffer from spherical aberration. If it could safely be assumed that under intense retinal illumination the cones of the human eye migrate farther into the pigment-layer we might expect high illumination to increase the resolving power of the retina, by decreasing the number of diffusion circles to which a single cone might be exposed at one time. The same result would

follow from suitable migration of the pigment itself with respect to the cones. Evidence as to the fact of migration in the human eye is lacking, however, so that the assumption is not safe.

The most important factor in vision, subject to the above limitations, seems to be the degree of stability of adaptation.

While the human eye can adapt itself almost equally well to a brightness between 1 and 100 millilamberts at different times, it cannot adapt itself to two widely different brightnesses at the same time. A considerable time is required for adaptation to a field which is considerably darker than one to which the eye has just been exposed. It is especially discomforting to have to fixate two unequally bright fields in rapid succession; some of my students have recently found that the productive output is seriously diminished thereby; and it is said that many industrial accidents have resulted from the necessity of the operator looking from a bright field into a shadow where moving machinery must be adjusted. This sort of disturbance of adaptation may be called temporal. A second sort, which may be called spatial, consists of having the central part of the field surrounded by brighter surfaces. If the area or the excessive brightness of the latter be sufficiently large, the eye is poorly adapted to the brightness of the former.

I shall here mention two everyday illustrations of the second sort of disturbance in order to fix the problem more thoroughly.

If a drafting-board be subjected to equal illumination from all directions, the brightness of its surface adjacent to a straight-edge laid upon it is just half the brightness of the surface a few millimeters away, since at the wall of the straight edge half of the illumination is occluded. The area of this shadow being very small as compared with the adjacent surface, the latter determines the state of adaptation. Ordinarily, as in tracing, the detail to be distinguished lies in the densest part of the shadow; and the draftsman, if the detail is small or faint, will find it quite invisible unless he casts a shadow with his body on the neighboring surface also. This

restores visibility, and relieves eyestrain at the expense of an awkward and sometimes uncomfortable bodily attitude. The work is greatly ameliorated by the use of a local lamp so placed that it is always in front of the T-square and to the left of the triangle, although the best results require additional precautions in the adjustment of the lamp.

If the material worked upon has a low reflectivity, such as that of dark cloth or rubber saturated in lamp-black, and if the tables and walls have much higher reflectivities than that of the work-material, or if the walls are filled with diffusing glass windows, the work-material, under 'general' illumination, may be seen in surroundings from six to 400 times its brightness. (This assertion is based on direct measurement.) If small and faint details have to be distinguished, as seams or thread, auxiliary local lighting is indispensable unless the undesirable alternative be adopted, of darkening the walls and surroundings.

Cobb and Geissler (2) and Cobb (3, 4, 5) have shown that the liminal separation and the liminal relative differences in brightness are least when the test-field and its surroundings are approximately of the same order of brightness; and that they increase slightly with large diminutions in the brightness of the surroundings below the point of equality, and increase very rapidly with increase in the brightness of surroundings. In (4) Cobb showed that the effect is due almost solely to the ratio of these brightnesses. Adams and Cobb (1) have more recently reduced the data in (4) to equational form, and have related them to certain postulates of nerve-response which harmonize with the 'all or nothing' law of Lucas.

In the year 1916–17 I found (17) that the time required for selective or discriminatory reaction to one of two stimuli was similarly affected by the illumination of the fovea with respect to that of the remainder of the retina; and I suspected that under the experimental conditions the latter test was more sensitive as well as more appropriate. The conditions of the present experiment did not demonstrate the superiority in sensitiveness which I expected to find; but the results otherwise justified its use.

Cobb's study (4) yielded data which in my opinion were sufficient to establish the principle in lighting which the present results indicate; but this view did not meet with instant acceptance in 1920 because the import of Cobb's results had not yet been fully assimilated by the engineering world. Also, they were based on rather few measurements; and limited to a test-field of foveal dimensions; and furthermore, the duration of exposure (3 seconds) being rather long, it seemed probable that a greater effect may have existed than the measurements had revealed.

It seemed advisable, therefore, first, to gather more data by the use of a field of foveal extent; and second, to relate such effects as might be found to the area of the central field, in order to show how large a primary field would render a given brightness-ratio tolerable. The second question is of importance because it is often impracticable to get a desirable distribution of illumination without interference with the operation done under it. Other duties prevented completion of the second part of the problem at the time, but the first was successfully completed.

The choice of the reaction-time method was dictated by considerations of convenience and appropriateness which, in view of certain current criticism, may require elucidation.

The Use of the Reaction-time Method

Subject to certain constant corrections determined by the particular chronometric apparatus which is employed, the time which intervenes between the presentation of one of a definite group of stimuli and the response of the muscles appropriate to that stimulus and to no other, is the time required for selective or discriminative response. This time is an index of the preparedness of the organism (1) to receive and (2) to respond selectively to, the several members of the group of stimuli.

The preparedness to receive the stimulus is affected by the limits of the sensitivity of the sense-organ and by its degree of adaptation at the time the stimulus happens to be applied. (In the special case of vision the 'sense-organs' are

the cones—or, at low brightnesses, the rods—of the retina on which the stimulus is imaged. The sensitivity of the sense-organ varies about a norm for each individual.)

The preparedness to receive the stimulus is also affected by the disposition of the 'accessory sensory apparatus' at the moment of presentation of the stimulus. (In the case of vision this apparatus consists mainly of (1) the muscles of the neck, which may have directed the face toward or away from the stimulus; (2) of the external muscles, which may have rotated the eye so that the stimulus is imaged on a more or less sensitive part of the retina, and which at the instant of exposure may be moving the eyeball or holding it in fixation; (3) the ciliary muscle, which may have disposed the lens so as to give a maximally sharp or a blurred image on the retina; and (4) of the eyelids, which at the instant of exposure may have been open or closed, in motion or still.

The degree of preparedness to respond to the stimulus is affected by the state of tonus, or partial activity, of the various sets of muscles which are to be 'selected' for response. This condition is influenced by fatigue and practice; but chiefly by the activity of circular neural arcs (originating and terminating within the given muscle) which has been initiated by the primary response made to the warning signal.

The first group of 'factors of preparedness' which I have mentioned may be called 'sensory'; the second and third 'attentional' or 'volitional,' although no mythological entities such as 'attention' or 'will' are postulated.

If the 'attentional' and 'volitional' factors could be perfectly stabilized, the selective reaction-time would be determined by the 'sensory' factors; and *vice versa*. If both operate at once the group whose influence is the stronger at the time will dominate.

Perfect stability cannot, of course, be secured; but the variability of a given group can be greatly diminished, the degree which is attainable being dependent on the professional skill and integrity of the experimenter in arranging the situation, and the degree of coöperation lent by the reactor in the following of his instructions. If this reduction of

variability be sufficient, the measured times are useful indexes, though not measures, of the influence of the experimental variable on the remaining group of factors.

The use of reaction-times in general has been recently disparaged (15), pp. 95–100, as if the lack of separability of extraneous factors under which it suffers is not a characteristic of other useful tests also. Such, of course, is not the case. In any situation to which an intact organism responds, the response is made, not to a single isolated stimulus, but probably to all the stimuli which are playing on the sense-organs at the time. Among the most numerous and important of these stimuli are *muscular contractions*—partial or complete—which excite the sense-organs ('muscle-spindles') imbedded among the muscle-cells. Therefore, in no situation can the attentional and volitional factors be isolated from the sensory ones.

The order of magnitude of a threshold, for example, is influenced by the subject's instructions, by his intentions regarding their execution, by his varying interest in the experiment, by his emotional attitude toward the experimenter, by his 'fear of being tested,' by his tendency to avoid mistakes at any cost or his willingness to hazard a guess, and by the manner in which his judgment is to be expressed. These sources of variability do not vitiate the use of threshold-measurements; they do call for maximal control, and for ample consideration of the degree of control which is secured.

While a chronoscope in the hands of novices may give misleading results, the measurement of reaction-times under controlled conditions is a good means of studying the appropriateness of the situations which are to be varied. For some purposes it is the best method available.

The Present Experiment

The present experiment can be briefly described. A circular test-field was prepared, whose diameter subtended approximately 3° at the eye; whose brightness (hereafter indicated by the symbol B_F) was sensibly uniform over the entire extent, and kept practically constant at 2.6 milli-

lamberts; and whose color is approximated by that of a black body at a temperature of 2400° K. In the center of this field was placed a visible dot to serve as a fixation-mark; on the horizontal diameter of the field and with their centers 45′ from its center, were projected, by means of two miniature stereopticons, two circles, whose diameters subtended 24′ at the eye of the observer. The removal of one or the other of these spots constituted the stimulus to reaction. In the preliminary work these spots had different brightnesses; but in the stages whose results are used their brightness was 0.09 millilambert; so that they constituted additions of 3.5 per cent. to the brightness of the field on which they were projected. The field was viewed monocularly; the unused eye being covered by a piece of ground glass invisible to the other. The remainder of the field of vision was filled by the interior surfaces of a polyhedral chamber described elsewhere (2), (5), (17). These surfaces were painted 'flat white' and were diffusely illuminated from invisible sources. Their brightnesses (B_S) were sensibly uniform with respect to each other but variable with respect to that of the test-field (B_F) and *they were the most effective determinants of the level of adaptation.* The variable factor tested was the ratio B_S/B_F.

Preparation of the Stimuli

The details of the manner of preparation and presentation of the stimuli may be understood from inspection of Fig. 6 and its legends. The lamp L_5 controlled the brightness of the test-field (B_F). This lamp was mounted behind a milk-glass screen on a movable carriage resting on an optical bench whose axis was normal to the test-field and coincident with its center. (This arrangement necessitated the field being viewed somewhat obliquely and rendered the projected stimuli slightly elliptical rather than circular. The field itself was truly circular, being limited by an opening in the box.) The surface of the field was quite matt, but the use of the milk glass was considered advisable. Lamps L_6, L_7 and L_8 controlled the brightness (B_S) of the interior of the box. This brightness was varied by variation of the number of the

lamps employed and further by the use of limiting diaphragms of different areas. All these lamps were operated on a generator. The circuit unfortunately was variable, and the lamps had to be controlled by the manipulation of the rheostats R_1 and R_2, used with the voltmeter *VM*. The latter was read simultaneously with the presentation of the stimulus,

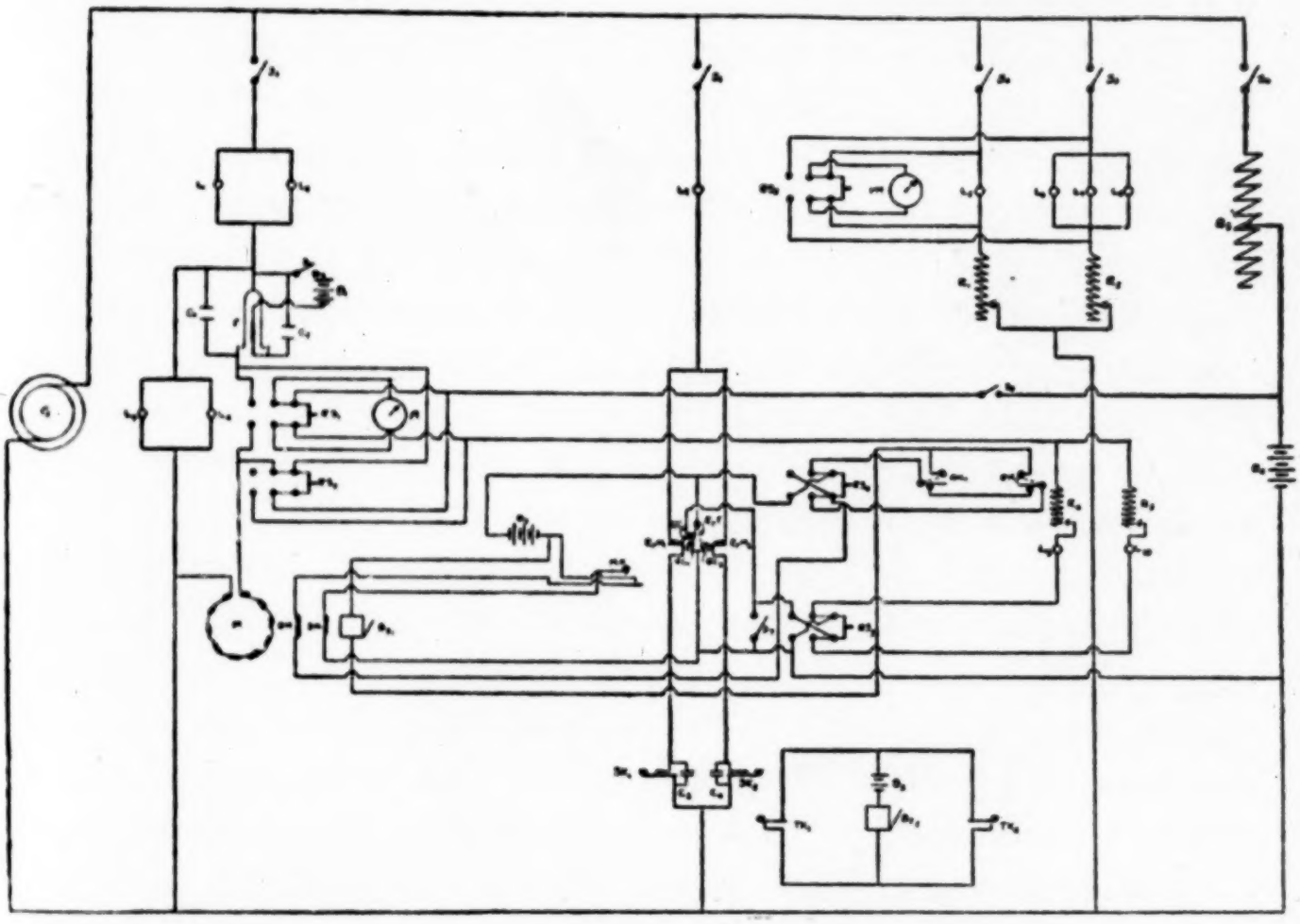

FIG. 6

EXPLANATION OF FIG. 6

G, 115 volt d.c. generator; B_1, B_2, B_3, dry cell batteries; L_1, L_2, L_3, L_4, lamps operated as resistances; C_1, C_2, C_3, C_4, 1 MF condensers; *F*, 50 v/s tuning-fork interrupter; *M*, motor of Dunlap chronoscope;

L_5, Mazda C lamp used to illuminate background; L_6, L_7, L_8, similar lamps used to illuminate the interior of the integrating polyhedral box; L_9, L_{10}, special flashing lamps used in stereopticons to form the removable stimulus-spots;

RS_3 and RS_4, commutating switches, operated by a single handle; RyM_1 and RyM_2, magnets, and *RyT*, tongue, of Dunlap relay; RC_1, RC_2 and RC_3, contacts at which circuits through lamps and chronoscope were made and broken by *RyT*;

MK, Dunlap master key; *RM* and *SM*, rotating and stationary magnets of Dunlap chronoscope; RK_1, *RK*., pneumatic reaction-keys; *Bz*, buzzer controlled by RS_4, registering erroneous reactions; SK_1 and SK_2, keys for actuation of relay;

Bz_2 and accessories, buzzer system for signalling;

R, variable resistance; *RS*, commutator switch; *S*, single knife switch; *VM*, voltmeter for control of lamps L_5, L_6, L_7, and L_8; *A*, ammeter for control of L_9 and L_{10} and *M*.

and if the reading was unsatisfactory the reaction was excluded. The *total range* of variation was not more than one volt above and below the normal. The mean variation is of course much smaller. The variability was undesirable because it rendered the differential brightness of the stimulus spot variable (its absolute brightness being constant), but this could not be avoided under our limitations of equipment.

The brightness of the stimulus spots was controlled by lamps L_9 and L_{10}. Their filaments were thin ribbons of tungsten, suspended in argon-filled bulbs. They were of the type designed by Lorenz, Worthing and Forsythe for use in military signalling. Their advantage for my purpose lay in the great speed with which the filaments cooled when the current was interrupted. It being necessary to control their current within 0.01 ampere, they were operated on a storage battery and controlled by means of sensitive rheostats and an ammeter. By means of the switch RS_5 in the diagram the current through either of these lamps could be returned directly to its source, while the current through the other was returned through two mercury-cup contacts RC_1 and RC_2, which were bridged by a dental gold wire attached to the movable tongue of a Dunlap relay.[1] The resistances of the direct and indirect paths were equalized by the inclusion of a short german-silver wire, not shown in the figure, in the direct path.[2] When the connection at RC_1 was broken the

[1] As may occur to the reader, the levels in these cups had to be adjusted so that the circuit through the stimulus lamp first broken at RC_1, then RC_2. This precaution offered no difficulty.

[2] This additional resistance was necessary, although the first ammeter which I used did not indicate the fact. The inclusion in the circuit through the stimulus-lamp of 1 cm. of heavy soft iron and 3 cm. of No. 18 gold wire reduced the luminous intensity of the stimulus-lamp and created a difference in brightness between the two spots which at the time escaped detection. After a considerable time I noted that one reactor (R) was giving a number of reactions, apparently not belonging to the main distribution, being grouped about 0″.16 with a range between 0″.13 and 0″.20. I suspected that they were really simple reactions, whose correctness was made possible by a sound basis of preferential expectation, which, however, was not often effective. In the search for clues I tried a more sensitive ammeter and found that the additional resistance of the relay was detectible. The resulting difference in brightness was practically at my own threshold, which is unusually low. I then required the several subjects to judge the difference; the most successful was R, who judged correctly in 114 out of 200 opportunities. The deviation of 7 per cent. from equality is twice as

lamp in series with it was extinguished and the stimulus-spot which it formed disappeared. The total time required for the brightness of the spot to diminish to an infraliminal value was about 0″.005. This time is a constant error of measurement, which is eliminated in differential comparison, and which is practically negligible in proportional comparison; therefore I did not subtract it from the measurements.

The time required for reaction was measured directly in units of 0″.01 by means of a Dunlap chronoscope. Reaction was made with one of two Dunlap pneumatic keys RK_1 and RK_2. These keys were connected in series. They were equipped with additional contacts against which the movable tongue struck when actuated. By means of the switch RS_4, whose handle was common to RS_5, the key which corresponds to the hand not to be used in reaction was placed in the circuit with a buzzer Bz_1, to inform the experimenter of the error when that hand was used. The chronoscope was started by the opening of RC_1, which broke the circuit through the stationary magnet of the chronoscope until the relay-tongue reached RC_2. The chronoscope was stopped by the opening of either RK_1 or RK_2. Thus the time of reactions made with the wrong hand or with both hands could be measured, as well as the time of correct reactions. The treatment of the incorrect reaction-times will be explained later.

The reactor was seated comfortably in a dark room adjacent to that of the experimenter, in front of the polyhedral box, and provided with a mouthpiece which enabled the head to be quickly adjusted. He was instructed to look at any part of the interior of the box, or at the test-object, as he preferred, between stimulations, but was cautioned against staring. About 1″.5 before each stimulation a warning signal was given by means of a buzzer Bz_2. The reactors were instructed as follows:

"When the warning signal is given, look directly at the

large as its probable error. He asserted that he had never noticed the difference. After the introduction of the compensatory resistances these unusually quick reactions ceased to occur. I regarded the previous lack of balance as a source of uncertainty, and discarded 12,000 measurements made prior to the change, which otherwise I should have used.

dot in the center of the test-object. When a spot disappears, press the *corresponding* bulb, *as quickly as possible.*

"Occasionally the signal will be given and the chronoscope started without removal of either spot. *Do not react in such cases.*

"Reactions made with the wrong hand, or with both hands, or without removal of the spot cannot be used.

"In the long run each spot will disappear as often as the other, but in a single series either may disappear the oftener. Try your utmost to hold yourself in equal expectancy for both spots during the waiting period.

"Signal the operator whenever you make a reaction contrary to these instructions."

The control mentioned in the second paragraph of the instructions was used from five to ten times in each series of 100 reactions during the preliminary training, but was eliminated in the regular work except in the series given for practice alone. It was accomplished by the use of an additional switch S_7, which short-circuited the relay-bridge. A compensatory resistance was used in series with this switch. The control aided materially in discouraging a tendency to anticipatory reactions.

Preliminary Training

Each subject was required to react between 2,000 and 10,000 times before he was used in the main experiment. Training was considered satisfactory when incorrect and anticipatory reactions were made only sporadically, when the reactor appeared at ease in the experimental situation, and when the variability of the measurements appeared to have become stable.

The Reactors

Five reactors were employed, as follows:

O. An electrical engineer, aged 32. Wore no glasses. Naturally left-handed, but by training ambidextrous. His understanding and attitude toward test were ideal.

K. A woman, aged 27; high school education; considerably above the average in intelligence; private secretary to

head of medical research department. Somewhat self-conscious; habitually studies her own performance closely, and shows a meticulous tendency to avoid mistakes. Preliminary training exhibited disturbance in any situation which suggested that she was being tested. The emotional phase disappeared but, as the record shows, the cautious attitude persisted. She took to training readily and learned quickly, but her average times are longer than those of the other reactors under the most favorable conditions Prefers local lighting for work and study Wore no glasses.

S. A woman, aged 23; college graduate; technician in bacteriology and pathology, medical research laboratory; accustomed to sustained, close and accurate visual observation. Behavior generally characterized by adaptiveness and self-possession. Wore no glasses.

P. Physician, aged 35; especially trained in physiology and internal medicine and distinguished by his achievements in those fields. Vision slightly above normal when corrected by glasses, which he wore during experiment. Poor muscle balance; fixation easily disturbed; at the time, an occasional sufferer from migraine. Usually concerned about his performance in the test; somewhat unstable emotionally and easily distracted.

R. A man, aged 19; office clerk; grammar school training. Visual acuity not accurately measured but far above normal. Wore no glasses. Despite careful and frequent explanations he regarded the experiment as a test of his mind. In part of the preliminary work he said he understood his task to be to make all his reactions of the same length. Later he phrased his intentions in the same words as the instructions, but I doubt if the original trend was ever quite broken up. He was trained with difficulty and imperfect success to obey instructions as to expectancy and effort; some 10,000 preliminary reactions being necessary to overcome a persistent tendency to guess what stimulus was to be presented. Apparently this tendency, though greatly lessened, was never completely overcome. He wore no glasses.

An Incidental Preliminary Discovery

In the preliminary work reactor *S* showed an insignificant and inconsistent difference in time between conditions "*C*" ($B_s/B_r = 1.14$) and "*D*" ($B_s/B_r = 3.87$) in the last 400 reactions with each hand under each condition, equally distributed. She frequently complained in severe terms, however, of eyestrain and discomfort under the latter condition, saying, "Unless I strain my eyes I can hardly see." Further questioning indicated that the trouble was due to difficulty in maintaining satisfactory fixation. I then tried the expedient of reducing the brightness of the stimulus spots to half their former value, so that they now constituted additions in brightness of 3.5 per cent. instead of 7 per cent. to that of the rest of the test-field. The results are given in Table III. They show that the diminution in intensity of the stimulus produced (1) a consistent increase in time under the unfavorable condition *D*, and a *decrease* in time under the favorable condition *C*. The effect is probably larger than the comparison indicates, since the comparative tests were made successively, and there was a slight improvement with practice. The interpretation, however, is clear. Up to a certain point, the more unfavorable the conditions of observation the greater the degree of 'volitional attentive effort.' This increased effort was sufficient to compensate for the unfavorable effect of *D* with the more intense and therefore more favorable stimulus 1; and it also compensated for the less intense and less favorable stimulus 2 in the more favorable surroundings *C*; but not in the less favorable surroundings *D*.

The indication is definite: reaction-time is not necessarily shorter under the sensorially most appropriate conditions unless the difficulty of the task under those conditions is sufficiently great to elicit a maximal degree of attentiveness. It would doubtless have been better to vary the intensity of the stimulus-spots in accordance with the sensitivity of the several subjects; but this step was rendered impracticable by the limitations of the apparatus. The intensity which I selected appears to have been about right for reactors O and

TABLE III

SHOWING INFLUENCE OF ATTENTIONAL ATTITUDE ON REACTOR *S*

Ratio B_S/B_F	Symbol	Left Hand							Right Hand						
		Stimulus 1		Stimulus 2					Stimulus1		Stimulus 2				
		M_1	E_{M_1}	M_2	E_{M_2}	$D, = M_1 - M_2$	E_D	D/E_D	M_1	E_{M_1}	M_2	E_{M_2}	$D, = M_1 - M_2$	E_D	D/E_D
3.87	*D*	27.68	.124	28.79	.181	− 1.11	.218	5.09	29.61	.141	30.93	.189	− 1.32	.236	5.60
1.14	*C*	27.57	.112	26.94	.145	0.63	.183	3.44	29.75	.130	28.61	.160	1.14	.203	5.62
$D, = M_D - M_C$	= 0.11			2.85					− 0.14		2.32				
E_D	= 0.194			0.23					0.192		0.248				
D/E_D	= 0.57			12.4					0.73		9.35				

Explanation: Relative intensity of Stimulus 1 = 0.07; relative intensity of Stimulus 2 = 0.035; *M*, average time in units of 0″.01 required for discriminatory reaction; under the conditions indicated by subscripts; *E*, the probable error of the determination indicated by subscript; *D*, the difference between the two averages given in each case. B_F, brightness of test-field; B_S, brightness of surrounding field.

This reactor was able to compensate, by an attentional strain, for the deleterious effect of the unfavorable distribution of brightnesses (*D*) with the favorable stimulus 1, but not with the less favorable stimulus 2. She likewise compensated for the deleterious effect of the unfavorable stimulus 2 under the favorable distribution of brightnesses (*C*) but not under the unfavorable distribution (*D*). The fact illustrates the necessity of a form of control which has often been overlooked.

K; but it probably was too high for *R* and possibly a little low for *S* and *P*.

The effect found for reactor S is characteristic for O and P also; but the comparison is less striking in their cases though scarcely less definite.

The Variable Conditions Tested

Six ratios of B_S/B_F were selected for final study. They will hereafter be referred to as follows:

Designation	B_S/B_F
A	0.015
B	0.34
C	1.14
D	3.87
E	6.43
F	9.45

Owing to contemplated changes in their duties, reactors P and S were tested only under conditions *B*, *C* and *D*. The other reactors were tested under all the conditions.

In all cases the order in which the conditions were tested within a daily series was arranged so as to minimize the influence of position within the series.

All the conditions in the list were tested in haphazard order before a test was repeated on any one. There is one apparent exception: in the case of reactor O, the results of one more series under conditions *B*, *C* and *E* are presented than under *A*, *D*, *F*. This is because one day's work under the latter group of conditions was spoiled by a stimulus-lamp becoming defective, and through oversight a part of the apparatus was dismantled before the test could be repeated. This reactor, however, was the most highly trained member of the group, having made 6,000 reactions in the work done preliminary to the main study. The practice-effect in his results which I have presented is negligible; so that the discrepancy in the number of cases has no other effect than of slightly increasing the probable errors of the differences between some of the compared averages

The right-left order of presentation was predetermined for the entire experiment by 200 tosses of 50 coins. These orders were made out in groups of 100, which were taken consecu-

tively. The slight discrepancies in the numbers of reactions with the right and left hands are due mainly to this fact, but partly because the number of errors made with the two hands was usually unequal for a given condition.

Each daily sitting consisted of three 'series' of 100 reactions each, one series being run under each of three ratios of B_S/B_F. Each 'series' was divided into four groups of 25 reactions each The first 'series' of each day was preceded, by 25 to 75 reactions for practice only, made under the first lighting condition to be tested in the sitting; the number of practice-reactions being determined by the subject's degree of training.

Each group of 25 reactions usually required from 3.5 to 4 minutes. An interval of about 2 to 5 minutes was allowed for relaxation and adjustment of apparatus between groups, and about eight minutes for adaptation and relaxation before each series.

The Results

The measurements yielded three criteria, namely, the average time required for discriminatory reaction, the variability of the reaction-times, and the relative number of errors.

Table IV. exhibits the average times (M), with the quadratic mean (σ) of the deviations from the averages, and the number of measurements (n) included in the averages, for each reactor, for each hand, and for each of the lighting conditions.

This table shows: first, that the average time, in every case, is shorter under condition *C*, which represents approximate equality of field and surroundings, than under any condition in which B_S/B_F exceeds this ratio. The increase in time is greater the farther this ratio is exceeded. Table V. exhibits the differences (D) between the average times under conditions *C* and *D* for each reactor, with each hand, together with the probable error (E_D) of the differences.

Table V. shows that under excessively bright surroundings of condition *D* there was a noteworthy increase in the average time required for reaction by subjects *O*, *K*, *P*, and *S*. This increase is between 0″.013 and 0″.038 in absolute magni-

TABLE IV

LEFT-HAND REACTIONS

Reactor	A			B			C			D			E			F		
	$\overline{BS} : BF = 0.015$			$\overline{BS} : BF = 0.34$			$BS : BF = 1.14$			$BS : BF = 3.87$			$BS : BF = 6.43$			$BS : BF = 9.45$		
	M	σ	*n*	*M*	σ	*n*	*M*	σ	*n*	*M*	σ	*n*	*M*	σ	*n*	*M*	σ	*n*
O.....	29.83	3.96	305	29.53	3.85	350	29.61	3.76	367	30.92	4.39	286	33.47	4.94	330	35.52	6.31	291
K.....	33.28	3.89	254	32.17	3.49	267	33.17	3.54	237	34.86	4.61	250	36.95	5.07	242	38.59	5.12	254
R.....	28.65	2.66	359	28.18	2.65	341	28.08	2.81	340	28.34	3.24	356	29.40	3.05	356	29.71	3.60	371
P.....				31.09	4.29	212	31.14	3.90	183	34.56	5.20	173						
S.....				27.53	3.96	254	26.94	3.43	253	28.79	4.17	243						

RIGHT-HAND REACTIONS

O.....	29.92	4.07	300	29.30	3.88	362	29.34	3.60	340	30.88	4.04	312	34.06	4.77	367	38.55	7.44	270
K.....	31.89	3.43	252	31.90	3.57	244	32.62	3.72	267	34.62	4.29	250	36.69	5.07	250	38.10	5.84	217
R.....	28.12	2.36	334	28.36	2.47	349	27.49	2.22	353	28.22	2.63	355	29.66	3.43	349	29.28	3.27	324
P.....				30.81	4.22	172	29.42	3.91	215	33.24	5.89	208						
S.....				28.60	4.36	246	28.61	3.74	250	30.93	4.47	256						

tude; between 4.4 and 13 per cent. with respect to the average time under the uniform distribution of brightnesses in condition *C*; and between 6 and 11.6 times its probable error.

TABLE V

DIFFERENCES (D) BETWEEN THE AVERAGE REACTION-TIMES (M) UNDER CONDITIONS C ($B_S''/B_F = 1.14$) AND D ($B_S/B_F = 3.87$) IN ABSOLUTE UNITS (0″.01), AND AS RATIOS TO THEIR PROBABLE ERRORS (E_D)

Reactor	Left Hand		Right Hand	
	D $(= M_D - M_C)$	$\frac{D}{E_D}$	D $(= M_D - M_C)$	$\frac{D}{E_D}$
O	1.31	5.95	1.54	6.14
K	1.69	6.76	2.05	8.55
R	0.26	0.68	0.83	6.75
P	1.85	7.99	2.32	9.41
S	3.42	10.3	3.82	11.6

On the assumption that the variability of these differences follows the Gaussian law—and I am under no illusions as to the character of the assumption—we might say that the probability of the smallest difference being as small as zero or as large as 0″.026 is of the order of 1 : 38,000; and the probability of the largest difference being as small as zero or as large as 0″.076 is of the order of 2×10^{-14}.

Judging from the fact that the tables of the probability integral function corresponding to x/E_x as given in text-books on the theory of measurement are usually not extended beyond a value of 6 for the argument E_x, we may conclude that the smallest difference which I have mentioned meets the accepted scientific standards of satisfactory approximation to certainty. If a higher standard is desired, one need only to increase the number of measurements, since if the variability is constant $E \propto 1/\sqrt{n}$. A great advantage of reaction-time measurement over other psychophysical procedures lies in the fact that a large number of determinations can be made in a relatively short time once the variable whose effect is doubtful is selected. In (15) it is objected that the method suffers from 'the difficulty of getting differences in result,

due to changes in the conditions tested, which are safely in excess of the mean error of the test.' We may take this occasion to remark that the reliability of a difference is properly judged by reference to its own probable error; and that this error depends on the number of measurements included in the compared averages as well as on the arithmetic or quadratic means of the deviations of their included measurements. The only important defect from which the reaction-time method suffers is one which is shared by every other method of studying the behavior of the intact human organism: namely of minimizing large and systematic variations in other effective agents than the one under study.

It will be noted that subject *R* shows a smaller effect than the four other reactors as between these two conditions. The difference, I suspect, is due in part to his greater visual sensitivity and in part to an attentional strain, which proved inadequate under conditions *E* and *F*.

In all cases, the effects of conditions *E* and *F* are greater than those of *D* and in the same direction. I shall therefore omit detailed discussion here.

To return to Table IV: it will be noted, secondly, that when the brightness of the surrounding field is made lower than that of the test-field, the effect on discriminative reaction-time is inconsistent. For all the subjects the effects are absolutely and relatively small and they vary in direction in individual cases. The general trend is toward a slightly lengthened reaction time for a relatively low brightness of surroundings.

If we next examine the distribution of the standard deviations (σ) we find more variability but a similar trend. In every case the magnitude of σ increases as the brightness of the surrounding field is increased above that of the test-field, and in every case but that of subject *R* the increase is greater the greater the excess of the ratio B_S/B_T above unity.

As this ratio becomes less than unity the standard deviation changes in different directions for different subjects. The character of the changes should be noted from close study of the table, for it clearly suggests definite changes in attentional attitude in certain cases.

In accumulating these measurements the subjects occasionally reacted with the inappropriate hand or with both hands. These times were measured but the apparatus did not indicate which hand had stopped the chronoscope. I regarded such reactions as 'overt errors' and weighted them doubly in computing the total number made, assuming that for each time the patient 'guessed' wrongly he also guessed correctly. Besides these measurements, certain others were excluded: viz., those which were recorded as correct but which the reactor signalled as improperly made; and also a few which failed to satisfy Chauvenet's criterion. The errors belonging to these two classes were weighted singly. Calling the total number of reactions made under a given condition by a given subject, with both hands, F, and the total number of correct reactions f ($= F$ less the sum of the weighted errors), the 'coefficient of accuracy,' a, would be expressed by f/F. Table VI. shows the coefficients of accuracy for each reactor according to the several lighting conditions.

TABLE VI

COEFFICIENTS OF ACCURACY $a = f/F$

Condition	Reactor				
	O	*K*	*R*	*P*	*S*
A	.945	.973	.960	—	—
B	.912	.987	.978	.944	.995
C	.946	.981	.963	.964	.981
D	.907	.971	.960	.931	.982
E	.922	.910	.935	—	—
F	.796	.896	.887	—	—

This table shows that the two ratios of B_S/B_F under which the longest times for reaction were required were also the conditions under which the largest percentages of error occurred. In the case of reactor *O*, the unduly large proportion of errors under condition *B* is evidently due to an additional effort, which was effective in preventing the average time from lengthening. Most of the errors occurred when the right-hand stimulus was presented. The reactor frequently

reported that he had first reacted with the right hand and had noted that the left hand twitched an instant later. However, I weighted such errors doubly.

Measure of Integral Performance

From the foregoing it appears that the average performance P of our subjects under any lighting condition may be evaluated by three criteria: namely,

(1) *Speed*, measured by the reciprocal of the mean reaction-time, $1/T$.

(2) *Constancy*, which may be measured absolutely by the conventional

$$h = \frac{0.707\sqrt{n}}{\sigma},$$

or relatively by

$$\frac{\sqrt{n}}{\sigma},$$

(3) *Accuracy*, which may be measured by f/F, in which f represents the number of correct reactions, and F the total number of reactions.

If any two of these factors were constant the performance would be measured by the other one. If all three vary, performance may be appropriately measured by their geometric mean or by their product. In this case I chose the product because it permits of a rough comparison between my results and Cobb's without necessitating additional computation. There is nothing arbitrary about this measure except the exponential weights to be assigned to the different factors. I elected to weight them equally, since the character of the data did not justify a more elaborate procedure.

This treatment had to be modified slightly because of the necessity of combining the performances with the two hands, which had separate measures of speed and constancy but an undistributed coefficient of accuracy.

The procedure adopted was as follows: Each average of time (T) was weighted according to the square root of the number of measurements ($\sqrt{n}$) which it represented; the reciprocal of this weighted mean was then multiplied by the

index of constancy, $\sqrt{n}/\sigma$. These products for the left and right hands were averaged and their mean multiplied by the coefficient of accuracy. (This treatment, applied to these data, gives results which do not differ appreciably from the results of a theoretically more defensible method in the most extreme case. Were the data less nearly homogeneous it might be open to criticism.) The final absolute measure of performance P' under any lighting condition was therefore

$$P' = \frac{1}{2}\left(\frac{1}{T_L \cdot \sigma_L} + \frac{1}{T_R \cdot \sigma_R}\right) \cdot a$$

Since our main interest does not lie in the exhibition of individual differences in absolute performance but in the effect of the ratio B_S/B_F on the performance of the individuals and of the group as a whole, I have taken performance (P_0) under condition C, in which B_S/B_F is nearly unity, as a standard of reference for each individual, and have expressed his performances under the other conditions as fractions thereof. That is,

$$\text{Relative performance} = P'/P_0'.$$

The results of this treatment are shown graphically in Fig. 7.

A comparison between these results and Cobb's (4) may not be out of place here. In Cobb's work the stimulus was a difference in brightness between the right and left halves of the test-field. At equality of brightness of field and surroundings the observer could compare each half, not only with the other half, but also with the adjacent surroundings. There is evidence that two of the three observers (C and M) utilized this advantage to some degree. I was the third observer and used a different method: namely, that of fixating the center of the field and rendering my judgment on its appearance at the instant of exposure of the stimulus, without continuing my attempt through the remainder of the 3 seconds of exposure in case the field appeared uniform.

In the present experiment this advantage of ($B_S/B_F = 1$) did not exist, as the stimulus-spots did not appear immediately

alongside the surrounding field. I therefore consider my own task in Cobb's experiment as most nearly representing that of my present reactors. In Fig. 7 I have shown my own performance in Cobb's experiment by oblique crosses, using as the measure of performance the reciprocal of the relative differential threshold for brightness. If the mean performances of all of Cobb's subjects were thus expressed, the data

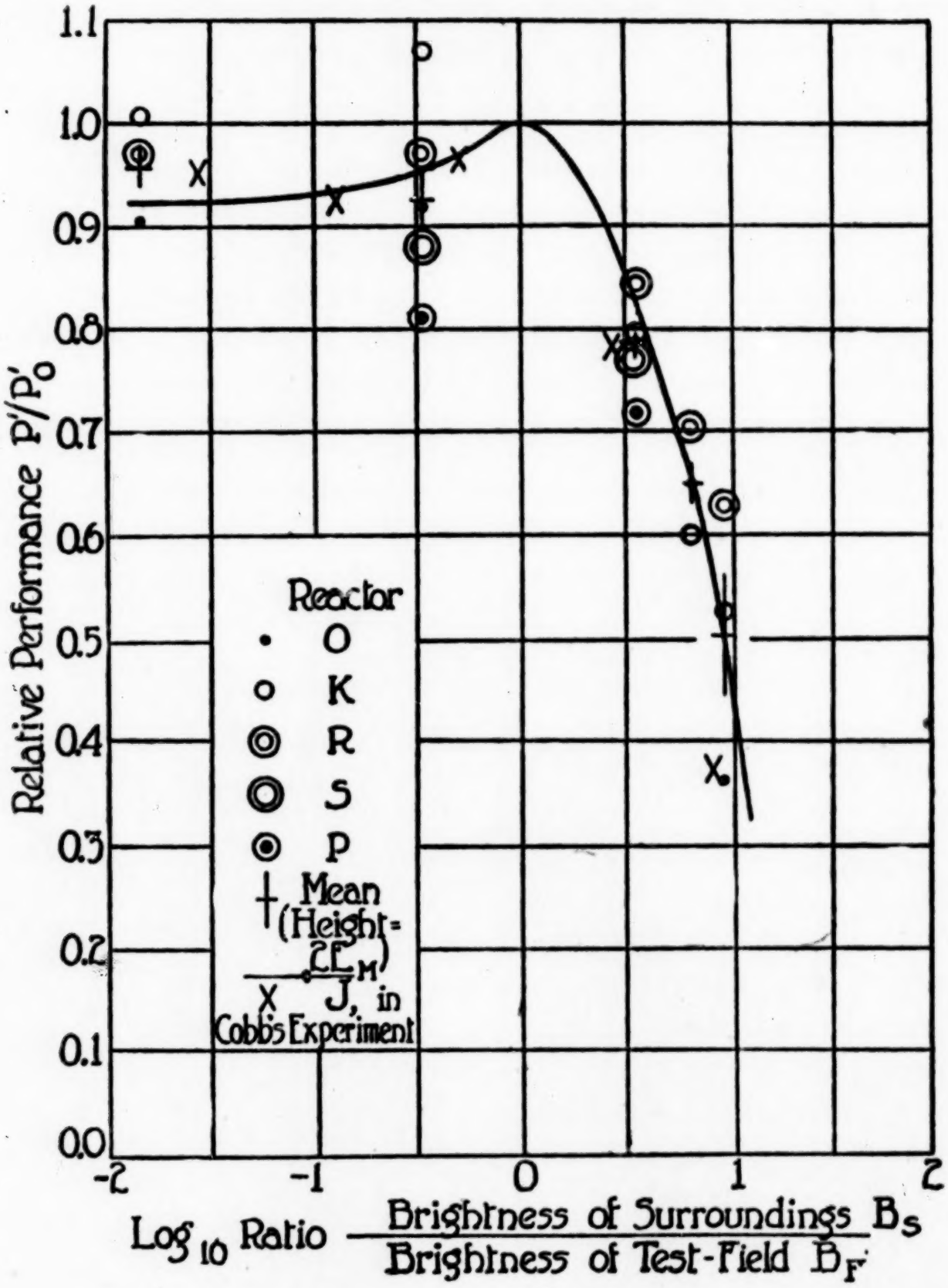

FIG. 7

could be plotted to a curve corresponding to the general equation of mine, but the fit at the abscissal value representing $B_s/B_r = 1$ would not be close. The reader may test this by experiment with the exhibit of Adams and Cobb (1), p. 44.

While the graphic comparison which I have presented may appear arbitrary under superficial inspection alone, it can be easily proved that whatever measure of average performance be adopted for either test the *correlation-ratio* will show a close correspondence between the two 'most probable' indexes, taken point for point along the abscissal scale. This fact indicates clearly that the two complexes of variables subjected to the respective tests shared their more important elements in common.

Discussion

The present results indicate what was suspected at the outset: namely, that the distribution of brightnesses is a capital factor in vision, in comparison with which, according to the work of other authors, the absolute level of mean brightness (so long as a certain minimum is exceeded) is of secondary importance. Practically, the results suggest that the use of auxiliary local lighting, despite its engineering difficulties, is necessitated in many operations in factory and office, in order to secure a tolerable distribution of brightnesses. The location of shadows, or the proper direction of specularly reflected light, may render a given visual task easy or enormously difficult. The proper lighting in a given case requires an analysis of the operation, first, from the standpoint of the visual task, and secondly, from the standpoint of disturbance of the workers' reaction-habits, as to posture, location of tools, movements, etc. This analysis, by whomsoever made, is a psychological problem, and must be made by methods appropriate thereto. It will reveal that in many operations the demands made upon vision are relatively scanty and that only rough considerations need to be taken into account. In less numerous cases, which, however, may be of extreme hygienic and economic importance, the matter of proper distribution must be taken up in detail. In all

cases the engineer's designs should be based on the results of such a study.

Acknowledgment

This work was made possible of successful completion by the generous coöperation of a number of friends: especially of Mr. G. W. Perk, late director of engineering, Mr. B. H. Clingerman, consulting engineer, Mr. H. M. Olds and Mr. A. R. Holden, electrical engineers, and Dr. Roy G. Pearce and Dr. Donald Lowe, of the department of health, of the B. F. Goodrich Company. It was greatly facilitated by the loan of certain apparatus by the research department of the National Lamp works of the General Electric Company, through Dr. E. P. Hyde and Dr. A. G. Worthing, and by the engineering department through Mr. Earl H. Anderson. Dr. P. W. Cobb and Dr. E. Q. Adams of the Nela Research Laboratories, and Professors W. S. Foster and K. S. Lashley of the University of Minnesota favored me with a reading of the manuscript and with certain apt criticisms thereof.

Summary

1. The factors in vision of absolute mean brightness of the test field and of the distribution of brightnesses over the whole field are compared and discussed.

2. The results of other students on the effect of absolute brightness on liminal differences in relative brightness, liminal angular separations of discriminands, and liminal durations of exposure are compiled, reduced to common bases and discussed.

3. Three reactors were tested under six ratios of the brightness of the surroundings with respect to the brightness of the test-field varying between 0.015 and 9.45, and two reactors under three of these ratios.

The results yielded three criteria: the time required for selective reaction, the dispersion of the measurements and the relative number of errors; or expressed inversely, speed, constancy and accuracy.

4. By all criteria taken both separately and integrally the results showed that performance is progressively impaired by

increases in the ratio B_S/B_F above unity and that rather small excesses of this ratio produced an effect which is important as to magnitude and reliability. The effect of reducing the ratio B_S/B_F below unity is small though the reduction be great; the average tendency is toward a slight and progressive impairment of performance, but some individuals exhibit a contrary tendency.

5. The results indicate that the group of variables subjected to this test includes the more important members of the group tested by Cobb in his measurement of the influence of the distribution of brightnesses on the relative differential threshold, the correlation ratio being high regardless of the mode of expression.

6. Subject to the restricting fact that the importance of the surroundings is diminished as the area of the central field is increased, the results exhibit the capital importance of the proper distribution of brightnesses, which in its turn necessarily indicates an increased use of auxiliary local lighting.

7. It was incidentally found that the reaction-time method may fail to indicate the influence of a stimulus-variable unless the conditions are so arranged that a discriminatory response to the most favorable stimulus requires a high degree of 'voluntary' attentiveness. In a certain sense it is advisable to 'calibrate' the reactors before the main experiment can be properly staged.

References

1. Adams, E. Q., and Cobb, P. W. The Effect on Foveal Vision of Bright (and Dark) Surroundings (*v*). J. Exp. Psy., 1922, *5*, 39–45.
2. Cobb, P. W., and Geissler, L. R. The Effect on Foveal Vision of Bright Surroundings. *Psy. Rev.*, 1913, *20*, 425–447.
3. Cobb P. W. The Effect on Foveal Vision of Bright Surroundings. *Psy. Rev.*, 1914, *21*, 23–32.
4. Cobb, P. W. The Effect on Foveal Vision of Bright Surroundings (*iii*, *iv*). J. Exp. Psy., 1916, *1*, 419–425, 540–567.
5. Cobb, P. W. The Momentary Character of Ordinary Visual Stimuli. *Psychobiology*, 1920, *2*, 237–244.
6. Cobb, P. W., and Loring, M. W. A Method for Measuring Retinal Sensitivity. J. Exp. Psy., 1921, *4*, 175–197.
7. Cobb, P. W. Individual Variations in Retinal Sensitivity, and their Correlation with Ophthalmologic Findings. J. Exp. Psy., 1922, *5*, 227–246.
8. Cobb, P. W. The Relation between Field Brightness and the Speed of Retinal Impression. J. Exp. Psy., 1923, *6*, 138–160.

9. Dunlap, Knight. A New Measure of Visual Discrimination. *Psy. Rev.*, 1915, *22*, 28–35.
10. Durgin, W. A. Lighting for Production (Lecture). Abstract in *Elec. Rev.*, 1919, *74*, 348 f. and 449–451.
11. Elliott, E. L. Artificial Lighting in the Rubber Industry. *India Rubber World*, 1921, *63*, 239–241; 329–332; 412–416; 483–486; 640–642.
12. Elliott, E. L. In: Discussion of Ferree and Rand's paper, *Tr. I. E. S.*, 1922, 89–92.
13. Elliott, M.: Comparative Cognitive Reaction-times with Lights of Different Spectral Character and at Different Intensities of Illumination. *Am. J. Psy.*, 1922, *33*, 97–112.
14. Ferree, C. E., and Rand, G. The Effect of Variations in Intensity of Illumination on Functions of Importance to the Working Eye. *Tr. I. E. S.*, 1920, *15*, 769–801.
15. Ferree, C. E., and Rand, G. The Effect of Variation of Visual Angle, Intensity and Composition of Light on Important Ocular Functions. *Tr. I. E. S.*, 1922, *17*, 69–86; with discussion, 87–102.
16. Ferree, C. E., and Rand, G. The Effect of Intensity of Illumination on the Acuity of the Normal Eye and Eyes Slightly Defective as to Refraction. *Am. J. Psy.*, 1923, 34, 244–249. (Same data as in (*14*), with respect to acuity.)
17. Johnson, H. M. The Influence of the Distribution of Brightnesses over the Visual Field on the Time Required for Discriminative Responses to Visual Stimuli. *Psychobiology*, 1918, *1*, 459–494.
18. Johnson, H. M. Reaction-time Measurements. (General Review.) *Psy. Bull.* 1923, *20*, 562–589.
19. Koenig, A., und Brodhun, E. Experimentelle Untersuchungen ueber die psychophysische Fundamentalformel in Bezug auf den Gesichtssinn. *Sitzb. Ak. Wiss.*, Berlin, 1888, 917–931.
20. Koenig, A. Die Abhängigkeit der Sehschärfe von der Beleuchtungsintensität. *Sitzb. Akad. Wiss.*, Berlin, 1897, 559–575; und: Gesammelte Abhandlungen, 378–395.
21. Nutting, P. G. *Bull. Bu. Standards*, 1908, *5*, 285.

AN EXPERIMENT ON THE EFFECTS OF LOSS OF SLEEP

BY HERBERT R. LASLETT

Stanford University

Introduction

The purpose of the experiment reported in this paper was a study of the effects of the loss of sleep upon the ability to do mental work, particularly upon three levels of mental work. These levels varied from the simple physiological retention of rote memory for digits through a substitution test which was largely reducible to a reaction time experiment with practice to an analogies test which required a considerable amount of thought and judgment. Many tests during the last seventy-five years have shown that motor tests of mental fatigue are of little or no value. The tests which involve mental work that can be reduced almost entirely to habitual effort, such as addition or cancellation of single letters, have proven little more valuable. It seemed probable, then, that tests involving the higher mental levels should show the greatest effects from the loss of sleep and the lower levels should show lesser effects.

The qualitative effects of the loss of sleep on mental products have been reported by Roemer (1) and Aschaffenburg (2) as increasing the proportion of sound and other lower grade associations in the free association test without increasing the association time. Patrick and Gilbert (3) reported visual hallucinations in one of their subjects and Miss Smith (4) reported emotional instability as an accompaniment of the exhaustion from loss of sleep. From a quantitative standpoint the results of Patrick and Gilbert, Miss Smith, and Hermann and Robinson (5) show that the result of wakefulness over relatively long periods causes a marked deterioration in some functions while others show little or no modification.

Patrick and Gilbert report that their subjects were restored by 16 to 35 per cent. of the sleep (measured in hours) that they would have taken normally, explaining this by the probability that the sleep following the long period of wakefulness was of greater depth than usual. The subjects in this experiment regained the normal numbers of hours of sleep through extra amounts in the days following the experiment.

It was not possible to make tests on the days following the experiment for the purposes of making curves of recovery. Miss Smith and Hermann and Robinson report such curves but they are not in close agreement.

Subjects

The experiment was conducted in conjunction with the department of chemistry of Stanford University, and was made possible by the kindness of Dr. N. W. Rakestraw and by the following students who served as subjects and controls: H. Q. Acre, C. V. Barley, L. E. Detling, F. S. Evans, C. E. Hesthal, J. L. Newell, D. L. Robertson, J. O. Vaughn, and I. C. Wilson. These were all young men of good physical stamina and suffered no ill effects from the experiment.

Method

The experiment is an endeavor to measure the effects of the fatigue induced by approximately fifty hours of continued wakefulness. The waking period extended from the ordinary time of rising on Friday morning until eight o'clock on the following Sunday morning. The investigation was made up of two experiments, separated by three weeks in time and similar in all ways except in subjects. The first experiment was begun on April 4, 1923, and continued until April 15, 1923 (both dates inclusive). One test period of about fifteen minutes was used each day. This was held at 5 o'clock P.M. except on April 7 when the tests were given at 11 o'clock A.M. to meet the convenience of the subjects; on April 14 when they were given at eleven o'clock A.M. as the mid-point of the sleepless period; and on April 15 when they were given at 7:30 A.M. as the end of the experiment. The sleepless period began at the time of rising on April 13 and continued

until 8 A.M. April 15. April 14 and 15 are, therefore, the critical days.

The second experiment began April 26, 1923, and continued until May 6. One test period a day was used as before. The tests were given at the same time of day with the same material used on the same day of the experiment, that is, the same material was used on the fifth day of the second experiment as was used on the fifth day of the first experiment. The same variations in the time of giving the tests were made on April 28 as on April 7, and on May 5 and 6 as on April 14 and 15, and for the same reasons.

The tests used in the experiment were (1) rote memory for digits, (2) a substitution test, and (3) an analogies test. The test of rote memory was made up of a daily series of three numbers of twelve digits each. These were read by the experimenter to the subjects who wrote them down in the reverse order from that in which they were read. It was not expected that anyone would be able to give a series of twelve digits in reverse order correctly, but occasionally a subject was able to do this. The instructions given to the subjects were to "save what you can of these numbers by any method that you please." The position of a forgotten digit was to be marked with a dash by the subject. A sample series is 9 7 1 6 3 8 2 5 1 4 7 3.

The code or substitution test in which letters were converted into other letters from a code is similar in its general features to the geometric forms test of Woodworth and Wells, the substitution test of Lough, and the symbol-digit test of Pintner and many others. The code form was before the subjects during the test. A sample form is given:

B C E F I J S U Y Z (original)
L J I N Z R D P O A (code)

S E J F C	C J I B F	S B U Z F
C Z Y B I	C Y S E J	I Y J C E
S J U B F	U Z F B I	S J I C F
U F Z E J	I Y B Z J	U E Y Z B

etc.

One hundred letters to be transcoded were on each page. The time given for the work was two minutes and fifteen seconds. The score was the number of letters correctly transcoded.

The analogies test was made by dividing the two hundred analogies standardized by Pintner and Renshaw (6) into groups of equal values. The score consisted of the number of analogies correctly solved times their values. Each day's test was different in material, except that one analogy was used twice in the series. The time given for solving the analogies was one minute. Table I. gives the daily records of each subject in each test in the first experiment. The subjects who underwent the sleepless period are marked as 'Subjects.' The controls, who lived as usual during the experiment and who had their regular amounts of sleep, are 'Subject' (control).

Results

Table I

Subject I

Date	Rote Mem.	Code	Analogies
April 4	18	57	
" 5	26	73	
" 6	25	70	
" 7	29	74	
" 8	29	80	
" 9	28	83	32
" 10	33	78	40
" 11	33	86	45
" 12	34	85	55
" 13	34	92	67
" 14	33	94	57
" 15	33	76	16

Subject II

April 4	—	—	
" 5	—	65	
" 6	19	82	
" 7	32	86	
" 8	21	93	
" 9	22	88	52
" 10	23	84	52
" 11	30	89	75
" 12	27	92	49.5
" 13	33	99	67.5
" 14	28	99	64.5
" 15	34	89	25.5

SUBJECT III

April 4	23	58	
" 5	30	72	
" 6	22	74	
" 7	26	85	
" 8	25	93	
" 9	22	89	24
" 10	27	97	64
" 11	34	101	75
" 12	21	80	88
" 13	26	94	88.5
" 14	24	101	64.5
" 15	29	92	17.5

SUBJECT IV (CONTROL)

Date	Rote Mem.	Code	Analogies
April 4	21	42	
" 5	24	51	
" 6	15	57	
" 7	25	62	
" 8	13	64	
" 9	23	68	36
" 10	17	57	32
" 11	15	68	60
" 12	14	67	44
" 13	11	69	49
" 14	18	70	50
" 15	—	—	—

SUBJECT V (CONTROL)

April 4		81	
" 5		84	
" 6		86	
" 7		83	
" 8		79	
" 9		87	101
" 10		89	105.5
" 11		82	96.5
" 12		82	93.5
" 13		81	104
" 14		87	94.5
" 15		90	96

THE CHARTS FROM EXPERIMENT I

Chart I. gives the curves for the rote memory for digits test for all of the subjects in the first experiment.[1]

[1] Control subject IV. took the last test at a later hour than the other subjects: it is therefore not to be considered but is put in as a matter of interest.

Chart II. gives the curves for the code or substitution test for all of the subjects in the first experiment.

Chart III. gives the curves for the analogies test for all of the subjects in the first experiment.

The critical days are marked ×.

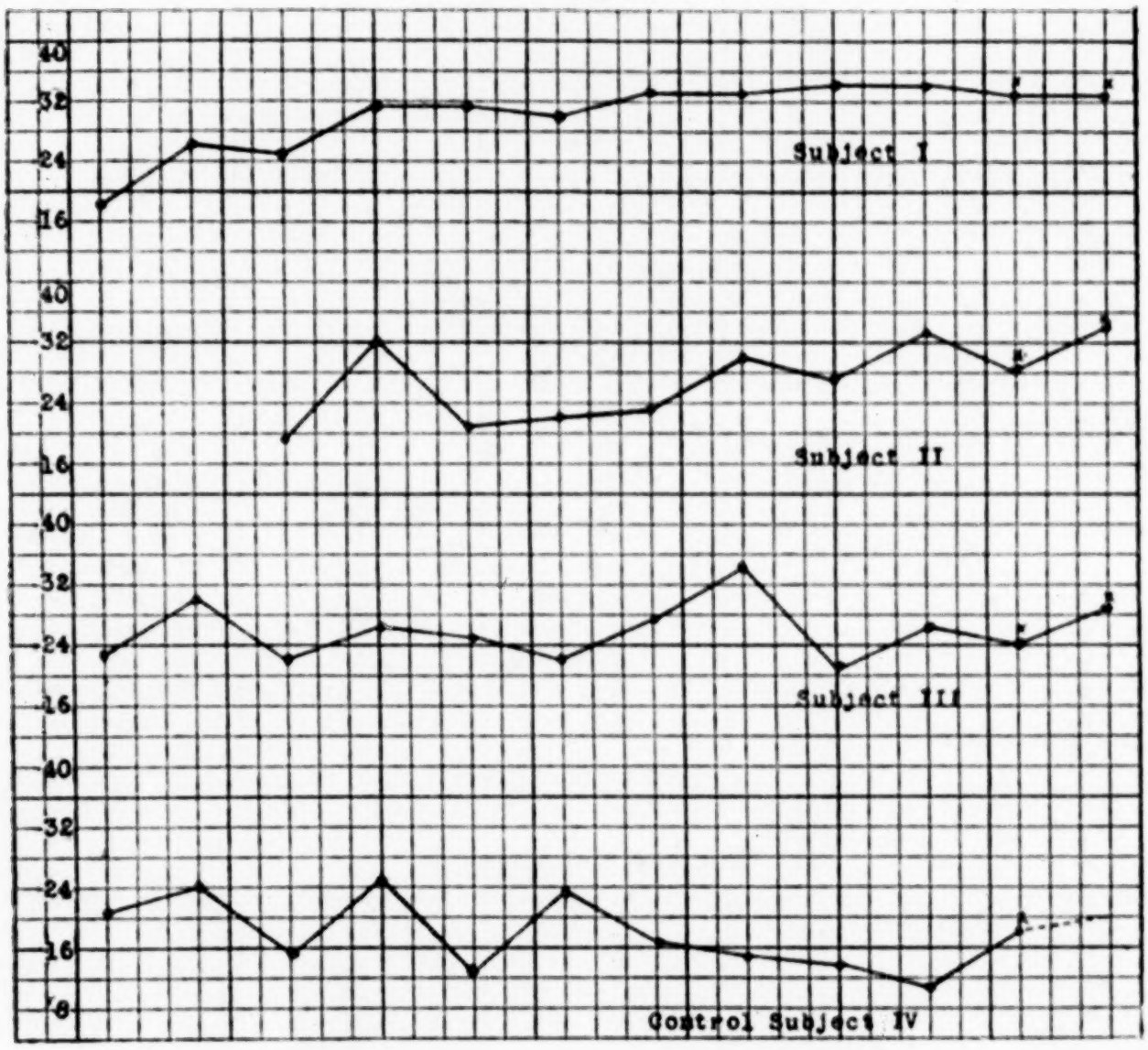

CHART I. Rote Memory for Digits

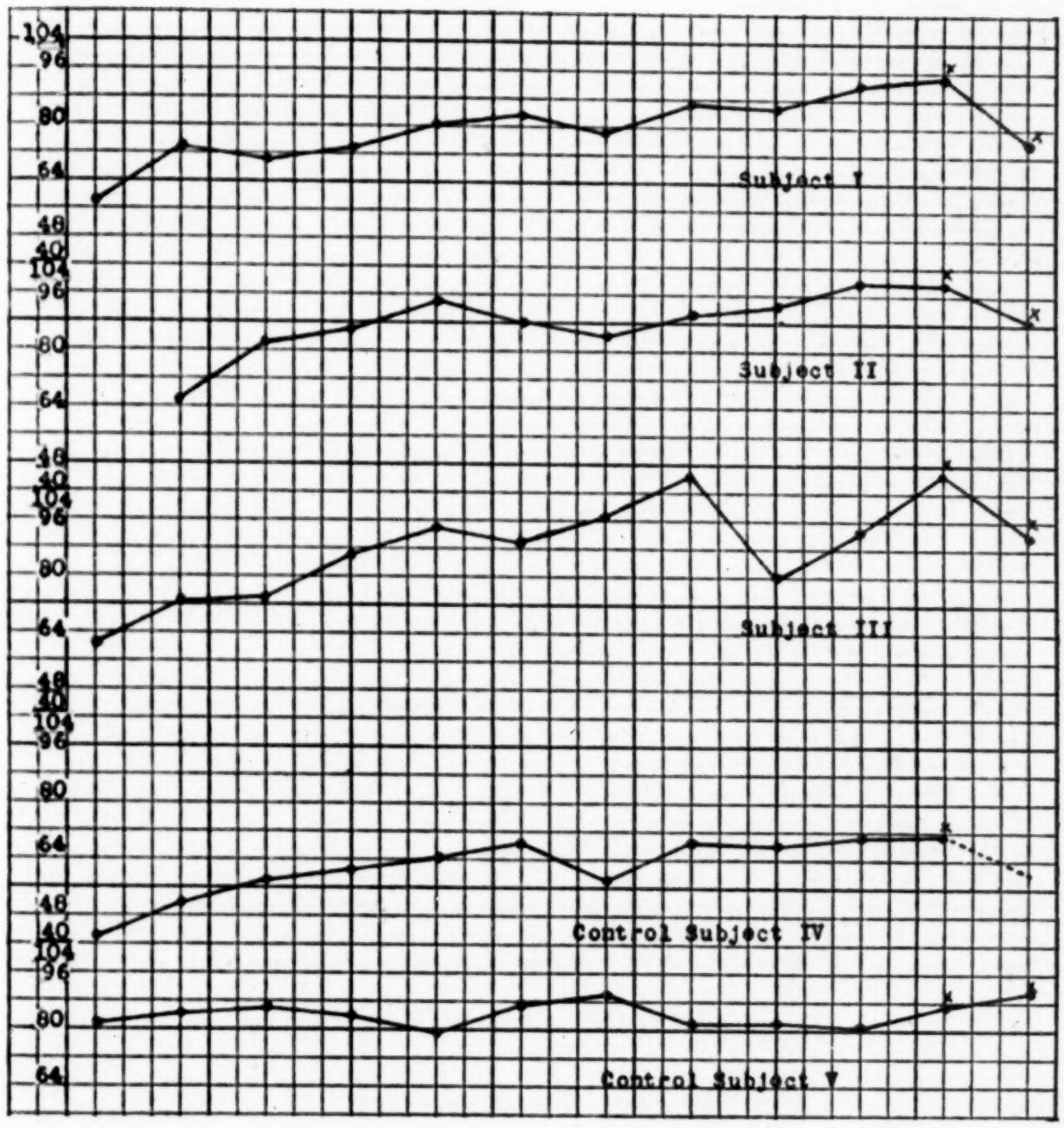

CHART II. Code (Substitution) Test

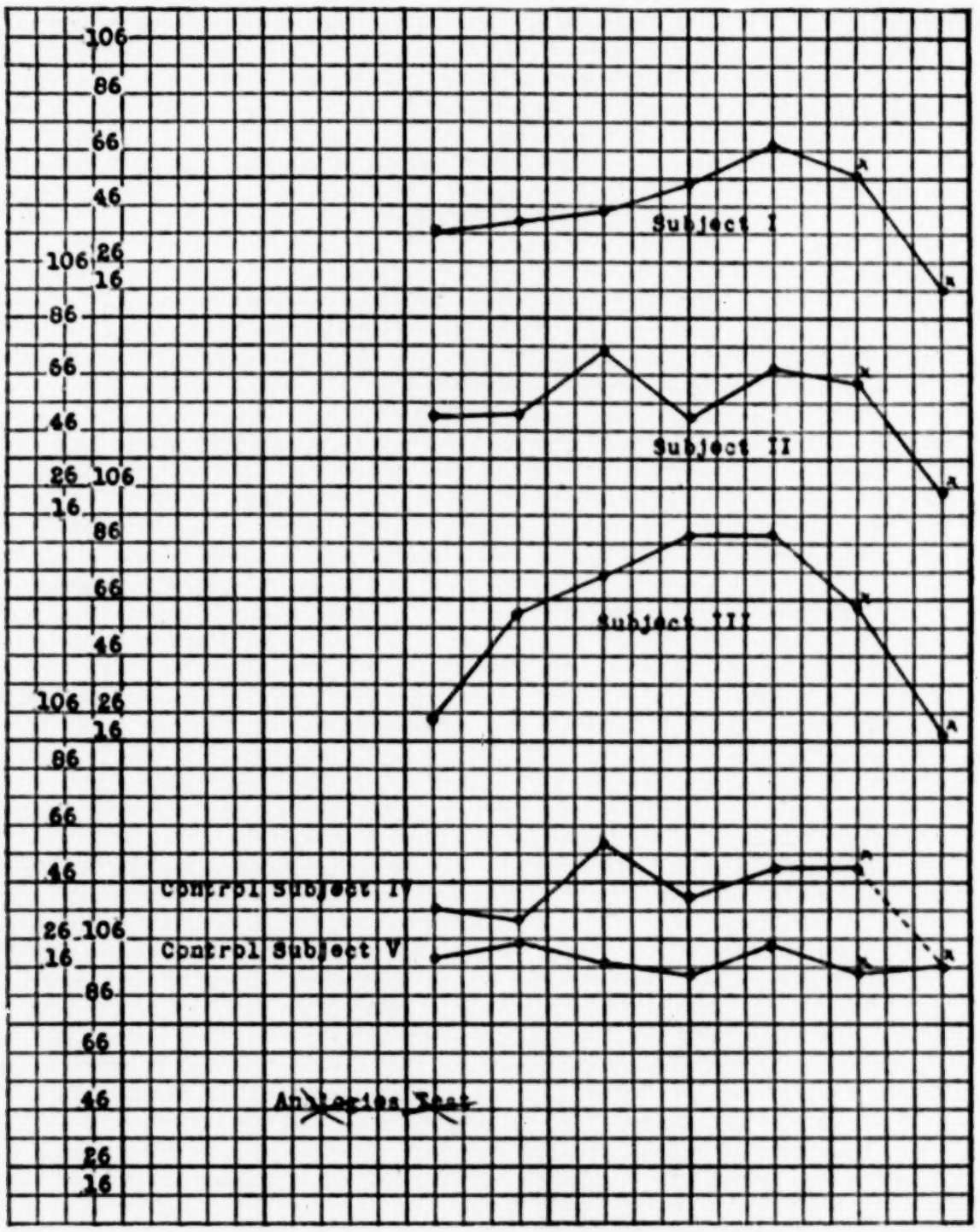

CHART III. Analogies Test

SECOND EXPERIMENT

Table II. gives the daily records of each subject in each test in the second experiment. The tests are the same as those in part of I. but the subjects are different.

TABLE II

SUBJECT VI

Date	Rote Mem.	Code	Analogies
April 26	10	59	
" 27	19	81	
" 28	—	89	
" 29	13	90	
" 30	12	110	24
May 1	12	91	36
" 2	17	100	45
" 3	25	97	33
" 4	20	93	36.5
" 5	11	90	35
" 6	13	97	24.5

SUBJECT VII

April 26	17	41	
" 27	15	44	
" 28	21	41	
" 29	16	51	
" 30	16	59	28
May 1	22	48	28
" 2	22	50	55
" 3	17	53	27.5
" 4	14	52	36.5
" 5	20	52	49.5
" 6	16	48	24

SUBJECT VIII (CONTROL)

April 26	17	51	
" 27	18	62	
" 28	23	62	
" 29	24	69	
" 30	25	72	32
May 1	22	64	60.5
" 2	22	63	50
" 3	13	72	44
" 4	23	74	80.5
" 5	19	79	71.5
" 6	24	72	62

SUBJECT IX (CONTROL)

Date	Rote Mem.	Code	Analogies
April 26	18	38	
" 27	23	45	
" 28	22	43	
" 29	22	58	
" 30	19	66	32
May 1	24	60	28
" 2	26	60	25
" 3	35	63	38.5
" 4	13	57	30
" 5	25	69	21
" 6	25	55	40.5

SUBJECT X (CONTROL)

April 26	9	58	
" 27	16	75	
" 28	20	72	
" 29	20	77	
" 30	27	83	60.5
May 1	16	66	57
" 2	14	86	80.5
" 3	22	60	55
" 4	23	71	67.5
" 5	21	82	42.5
" 6	17	66	52.5

THE CHARTS FROM EXPERIMENT II

Chart IV. gives the curves for rote memory for digits for all of the subjects in the second experiment.

Chart V. gives the curves for the code test for all of the subjects in the second experiment.

Chart VI. gives the curves for the analogies test for all of the subjects in the second experiment.

Since the results must be interpreted relatively, an examination of the curves will show the trend of the gross scores during the course of the experiments. The subjects (those who went without sleep) on the whole lost just a little during the critical period. The losses of the control subjects in the analogies test I am unable to explain, but I believe them to be individual and not due to the construction of the test.

An examination of the curves shows great irregularities

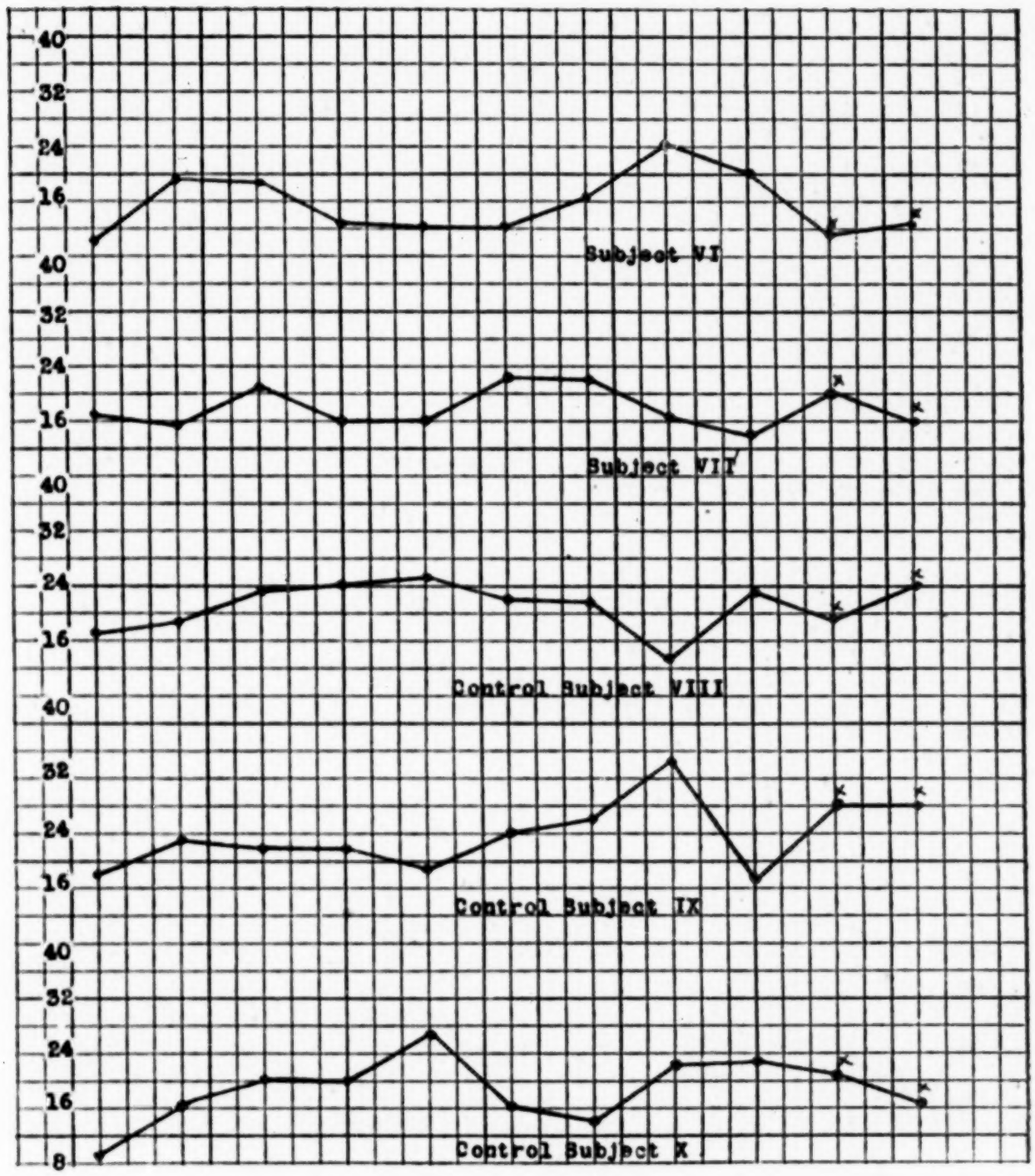

CHART IV. Rote Memory for Digits

from day to day. Some of them can be explained partially as due to sickness, colds, examinations just taken, and eagerness to finish the test in order to play tennis. Two of the subjects took one of the tests after a short, lively game of tennis, and made higher scores than usual, probably because of the increased bodily tonus. These influences should not have as great effects upon the test results as did the period of sleeplessness although they seem to have had. In the main, however, the irregularities are to be explained as daily fluctuations, the causes of which are unknown. Introspec-

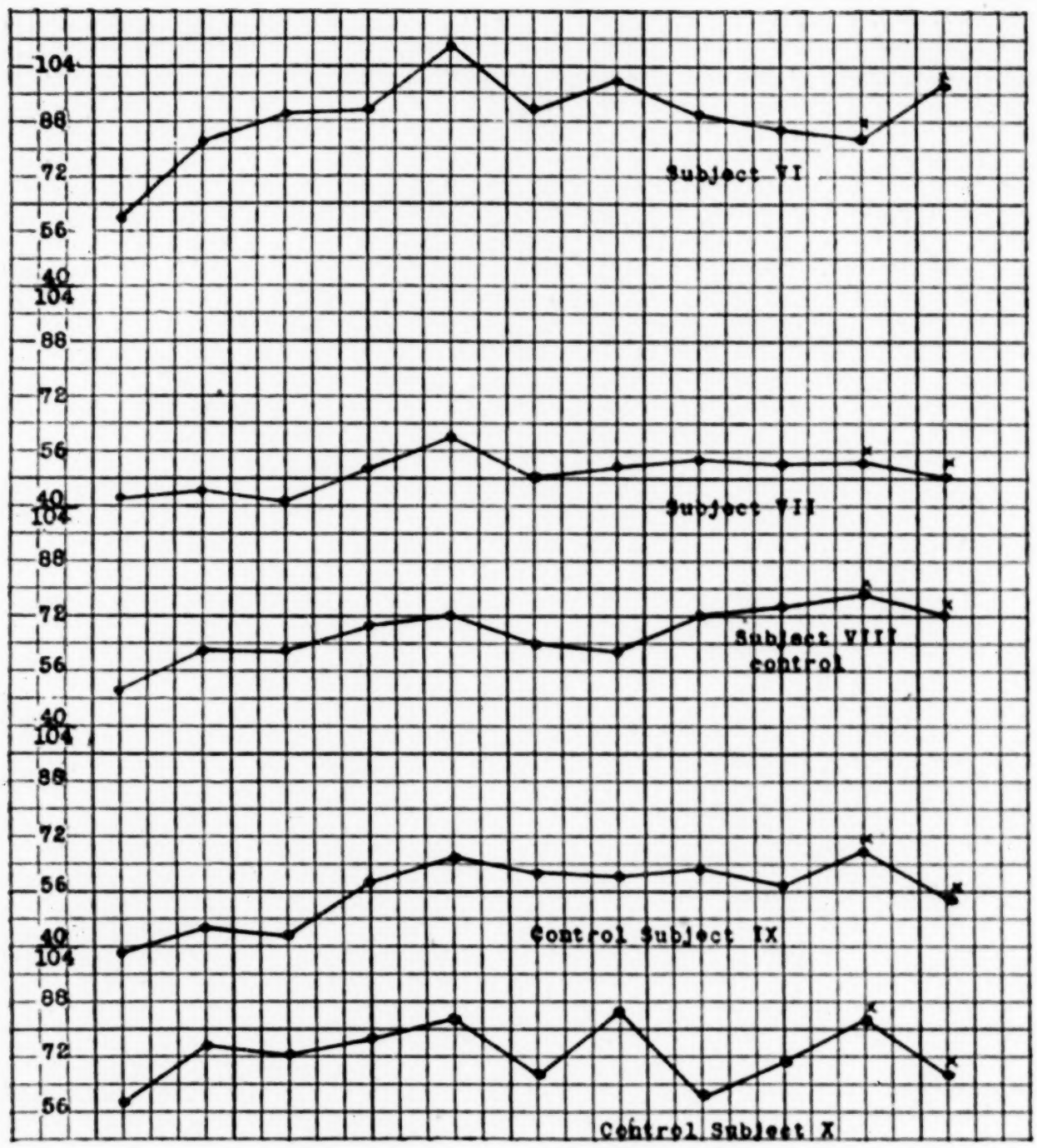

CHART V. Code Test

tions of the subjects and the controls on this point gave no explanations of the results.

The results are far from conclusive, perhaps even from definite. I believe, however, that they indicate a number of things.

1. Compensation by extra effort after one or two nights' loss of sleep is a matter of individual difference, reserve strength, and the length of the trial or individual test and its complexity or appeal to the higher mental centers.

2. The tests that are least subject to chance variation or to the influence of habituation or that are largely of a motor

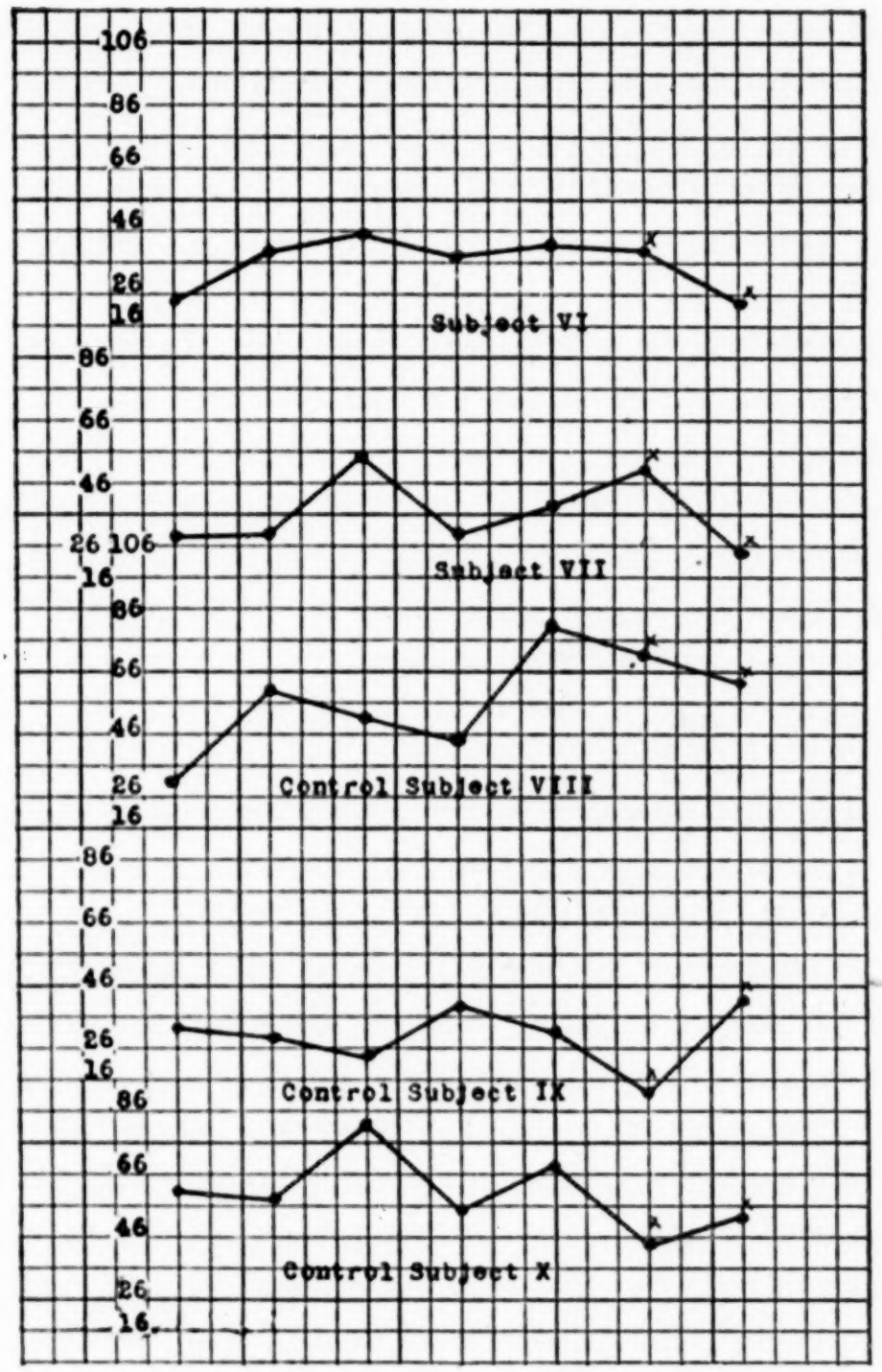

CHART VI. Analogies Test

nature are not as good tests of the effects of mental fatigue as those that involve the more complex mental processes.

3. Fifty consecutive hours of wakefulness are not enough to produce more than an indication of the results that would probably be obtained from a longer period of wakefulness.

4. Extraneous influences, such as colds, sickness, worry, eagerness for a holiday, trial of new methods of work may be of as great influence on the gross scores of these tests as short periods of sleeplessness.

5. Reserve strength, attitude towards one's ability to

withstand the effects of the loss of sleep, the novelty of the situation probably affect the scores greatly in individual cases.

6. Subjective estimate of the amount and quality of the work done in any trial is valueless. (The introspections are not given here but enough of them were taken to show this clearly.)

7. I found no 'spurt' after the slight fatigue of the loss of sleep for one night, thus failing to confirm the results of Miss Smith.

8. The subjects had a few momentary hallucinations during the course of the second night of sleeplessness, according to the reports of an observer, who stayed with them to see that they did not fall asleep.

Possible Criticisms of the Experiment

1. The period of sleeplessness was much too short.
2. The practice period was too short.
3. The tests themselves were too short.
4. A short period of work, such as multiplication or addition, should have preceded each day's trials in order to bring the subjects to the same degree of alertness.

Of the first two objections I may say that the periods of sleeplessness and of practice were as long as could be obtained under the conditions. In answer to the third, there is a great dearth of standardized material available for such experiments as these. The fourth criticism is hypothetical and must be subject to experiment, but the introspections of the subjects seem to show that part of the time of the trials was, in some cases, used to overcome the effects of inertia.

References

1. Roemer, E. Tätigkeiten 3rd Internationalen Congress für Psy. 1896.
2. Aschaffenburg, G. Experimentelle Studien über Associationen. *Psy. Arbeiten*, Vol. 1, 1895, and Vol. 2, 1897.
3. Patrick, G. T. W., and Gilbert, J. A. On the Effects of Loss of Sleep. Iowa University Studies. 1897.
4. Smith, May. A Contribution to the Study of Fatigue. *Brit. Jour. of Psy.*, Vol. 8, 1915–17.
5. Hermann, E. S., and Robinson, S. O. The Effects of Loss of Sleep. *Jr. Exp. Psy.*, Vol. 5, 1922.
6. Pintner, R., and Renshaw, S. A Standardizing and Weighting of Two Hundred Analogies. *Jr. App. Psy.*, Vol. 4, 1920.

VOCAL AND MANUAL MECHANISMS IN CHOICE REACTIONS

BY F. L. WELLS

Boston Psychopathic Hospital

These experiments were undertaken for the study of various choice reaction processes under cognitive settings. It was desired to compare them with choice reaction under affective settings, which will enter into subsequent discussions. They also have some bearing on behavioristic formulations. The 'choice reaction' is the regular name for the measurement of how long it takes to discriminate an experimental situation and act according to it. The setting may be to raise the right hand on seeing a white color; the left for black. This is among the least complicated of choice reactions. It involves only two choices of a quite elementary character. A choice reaction is necessarily differential; there must be as many different kinds of reaction movements as there are situations to be discriminated. Reacting with the different fingers has been tried, and reactions such as with the feet are possible but for various reasons these have not proved convenient techniques. Most choice reaction experiments make use of the right and left hands, and are thus confined to the discrimination of two situations only. In spoken language, on the other hand, one has a mechanism which is exquisitely differential, and can be applied to measuring the time of mental processes in indefinite number and complexity. With the voice, it is not generally practicable to record the nature of the response made, as can be done with the telegraph keys actuated by the hands or fingers. The response is usually understandable, however, and this is a practically sufficient check.

The simple voice key used in the present experiments was described in a previous communication (this JOURNAL, 1922, 5, 419–427). In the results there reported, loudness of

utterance made considerable difference to the action of the key. The initial sounds of the response words make less difference, but are another possible source of error. In the present experiments it was desired to compare a considerable number of mental processes, involving both hand and voice reactions. For a single experimental session of this kind, the usual procedure is to begin with a series of measurements of one type, then make the measurements of the next type, and so on until all series are gone through (*Cf.* the classical studies of Henmon, on the 'Time of Perception as a Measure of Differences in Sensation' *Archives of Psychology*, 1906, No. 8). On succeeding days, the order of the series is reversed, to guard against such intercurrent effects as are commonly known as practise or fatigue. Reversing the series is undoubtedly better than not doing so, but its effectiveness is modified by the consideration that warming up and fatigue do not operate continuously throughout an experimental session. Their effects are more prominent at definite parts of it, such as the beginning, or near the end. With these considerations in mind, it has been a frequent practise in this laboratory to take but one reaction of each kind at a time, until the entire range of reactions to be covered has been completed, and then repeat the process as many times as desired. If the experiments are to cover several sessions, the order of the processes is reversed from session to session. If, as is often the case, only one session of a sort is contemplated, the cycle is reversed in direction as soon as completed. A series of this kind is obviously harder work for both experimenter and subject. Against the first of these objections, and perhaps the second also, it can be urged that it is a more error-free type of experiment. With regard more specifically to the question of time, it may be noted that the actual time of observations is commonly a minor portion of the total time devoted to such an inquiry. A relatively great increase in the time devoted to observations does not proportionately affect the time devoted to an entire piece of work.

For the hand reactions in these experiments, two telegraph keys are used. The exposure apparatus is modeled on that

described in Benedict and Dodge, "Psychological Effects of Alcohol," pp. 94–5. The times are measured with a Leeds and Northrup ballistic galvanometer operating as a chronoscope. The sound stimulus used in the simple reactions observed is the click of a telegraph relay.

The present experiment involves thirteen different types of reactions. As explained above, the entire cycle of thirteen would be gone through and then repeated. The number of repetitions was ten, making 130 reactions for each session. A minimum time for such an experiment was fifty-five minutes. A subject could not be expected to remember from reaction to reaction what would be called for next. Verbal reminders were given at first, and throughout there was posted before him a schedule, giving the order of reactions and brief instructions for each. To reproduce this may perhaps be the most intelligible way of presenting them to the reader.

Item

1. Stimulus; black and white in 'light box.'[1]
   Press down both keys at 'ready' signal.
   REACT TO WHITE by releasing key on side at which WHITE is shown.
2. Same as (1) except
   REACT TO BLACK by releasing key on side at which BLACK is shown.
3. Black and white in drop screen.
   Press down both keys at 'ready' signal.
   REACT TO WHITE by releasing key on side at which WHITE is shown.
4. to 8. are same as (3) except
   REACT TO BLACK by releasing key on side at which BLACK is shown.
5. If white shows on RIGHT say 'Right.'
   If white shows on LEFT say 'Left,' *i.e.*,
   WATCH the WHITE—Say *Left* or *Right*.
6. If black shows on RIGHT say 'Right.'
   If black shows on LEFT say 'Left,' *i.e.*,
   WATCH the BLACK—Say *Right* or *Left*.
7. If black shows on RIGHT, say 'Black,'
   If white shows on RIGHT say 'White,' *i.e.*,
   WATCH the RIGHT—Say *Black* or *White*.
8. If black shows on LEFT say 'Black.'
   If white shows on LEFT say 'White,' *i.e.*,
   WATCH the LEFT—Say *White* or *Black*.
9. 9–13, Sound stimulus, click of relay.
   'BLACK' as simple vocal reaction.
10. 'WHITE' as simple vocal reaction.
11. 'RIGHT' as simple vocal reaction.

12. 'LEFT' as simple vocal reaction.
13. Simple hand reaction.
    Press down key at 'ready' signal.
    Release when click is heard.

There are in all fifty-two records of this experiment, comprising ten with four subjects, nine with one subject and also two and one with single subjects. These observations total 6,760. The presentation is based on the experiments with the five subjects first named, three women and two men. All these are professional psychologists, of varying experience.

The question raised in these remarks is how the vocal mechanism compares with the manual mechanism in executing the types of differential reaction called for. Thus, the experimental task in item 3 is essentially the same as in item 5 except that in item 3 the reaction to white is by raising a certain telegraph key, in item 5 by saying a certain word, 'white.' It is reasonable to question how these two processes compare in readiness. If one compares the figures as they come from the chronoscope, the vocal reaction averages about half again as long as the manual, but this result is subject to a correction. The voice key does not respond to speech so quickly or so regularly as does the telegraph key to the raising of the hand. The best available, though by no means a perfect, way to gauge this error is to compare the response words used when given as simple reactions, to simple reactions with the hand. This is the reason for items 9–13. When this is done, it is found that the vocal simple reactions average some 96 sigma longer than the manual reactions. There is but insignificant variation with the character of the initial sounds. This is probably because the subjects got into the way of barking the response words into the key, slurring initial consonants. If one subtracts from the chronoscope readings for the vocal choice reactions, each person's average excess of vocal simple reaction time over manual simple reaction time, the relations between items 3, 4 (manual) and 5, 6 and 7, 8 (vocal) are as follows:

[1] Items 1 and 2 were taken with a Dodge Mirror Tachistoscope, operating as an exposure apparatus. They do not figure in the present discussion.

Per cent. which the vocal times of items 5, 6 and 7, 8 are the manual times of items 3, 4:

Subjects	I	II	III	IV	V
Items 5, 6	113	154	122	154	182
Items 7, 8	103	105	111	126	127

All differences are still in favor of the manual reactions but they are seen to vary a great deal from subject to subject. Vocal reaction of this type is relatively disfavored with Subjects IV. and V., while for Subjects I. and III. the difference is smaller. With Subject II. it depends on the type of vocal reaction involved. It seems generally to take longer to record decisions of this level with the voice than with the hand. Subject V. had had more practise in choice reactions of the manual type, but this was not significantly the case with the other subjects.

As noted, the essential difference in these reactions is that in one case one raises the right or left hand, in the other one says 'black' or 'white.' In the first case one does the same thing, releasing a key, but with different parts of the body, *i.e.*, right and left hands. In the second case one uses in each event the same system, the vocal organs, but makes differential movements therewith. The general question is raised but not answered, as to the relative difficulty of responses differentiated in these two ways. Measurements of differential reactions with a single hand, moving in different ways, would be needed for this purpose. This matter is now being taken up.

Attention may be directed to the difference between items 5, 6 and 7, 8. Both are vocal choice reactions. But in the former one naturally observes both sides of the stimulus card, and selects the one on which black or white appears. In the latter, one naturally observes one half of the stimulus card, neglecting the other, and reacts according to the stimulus showing there. Both problems can legimately be classified as 'distinguishing black from white,' but the above methods of approach make a difference in the times observed. Items 5, 6 average about 127 per cent. of the length of items 7, 8. Their mean variation is 160 per cent. of items 7, 8. Evidently

the mental task is more difficult in items 5, 6. To some extent one might reproduce the conditions of items 7, 8 by a 'set' like the following, for example, in the case of item 5: The subject watches the right side of the stimulus card. Then if black comes down on that side, the response is 'left'; if white shows, the response is 'right.' An experimental set of this kind is different from that above described, and could hardly be expected to yield similar results. Yet the difference is not controllable save by introspection.

Manual reaction corresponding to the vocal items 7, 8 does not figure in the present observations. The conditions are, however, approximated in showing an entirely black or white stimulus card, having the right hand raised for white, the left for black. A number of such observations have entered into the experimental series utilized, but the subjects are not the same individuals as here concerned, and there is no direct comparison with exposures of the two stimuli in one experimental series. Yet it is evident that an expression like 'distinguishing black and white,' or any other members of a choice reaction series, is subject to a variety of meanings, according to the mental set, or the behavior processes involved in the differential reaction.

The physical conditions of these vocal reactions are such as to make them more variable than the manual, and this is the general situation. The average of all mean variations is 40 sigma. For the simple manual reactions, it is 18 sigma, averaging 11.5 per cent. of the reaction times themselves. Variability of this kind is often thought of as a measure of attention, and thus brought into relation with rather basic temperamental characteristics. The global figures for variability do not emphasize such differences in the subjects, though they patently exist. The extreme range is from 33 to 52 sigma. Scrutinizing the results more in detail, it seems that the more difficult processes reflect the differences better, as one might expect. The most difficult process of the experiments is that denoted by items 5, 6. It is also the most variable, its global m.v. being 80 sigma. The components of this figure also show the greatest range found, 53 to 118 sigma, and

these differences seem consonant with the degrees of attentional control brought to bear on the experiments as a whole. It is striking that in the simple manual reaction experiments, the range of variability is also great, ranging from 9 to 33 sigma (6.7 to 19.2 per cent. of individual averages). In part this is due to the excessive practise of one of the subjects in simple reaction work. The general facts represented are in accord with the data of items 5, 6. The average variability of these simple reactions is hardly more or less than that found in the other types of reaction. The only outstanding difference, that of items 5, 6, is owing to the relative mental difficulty, not to vocal or manual factors. In these experiments indeed, the simple vocal reactions show proportionately somewhat smaller variations than do the simple manual ones. The reason for this is, perhaps, that in the closest parallels which daily life offers to laboratory reaction experiments, it is vocal reactions rather than manual ones which are called for. The typical subject comes to the experiment with more practise in calling out something on a given signal, than in making a prescribed move with a telegraph key.

This opens the question of correlation between vocal and manual reaction times. Not enough subjects are represented to deal with a question of correlation as usually understood. On the other hand, the relative positions of the subjects in these functions are determined with fair adequacy. The question becomes whether individuals rapid in items 3, 4 are also rapid in items 5, 6 and 7, 8. Subject V. averages the next to shortest manual choice time (303 sigma), and much the longest vocal times (552, 384 sigma). This subject had relatively much practice in manual choice reactions, which may account for his short manual times, but makes it noteworthy that this practise did not carry over to the vocal functions. Subject IV., without special practise but with considerable manual dexterity, has much the shortest choice times, (250 sigma), but has no outstanding position in vocal reactions. The situation with the simple reaction times is essentially similar. It is probable that special manual ability carries over into the manual reaction times without affecting the

vocal processes and vice versa. This would much disturb any inherent relationship of the two. A further disturbing factor of serious possibilities is the dependence of the vocal response time on the loudness of the voice.

In this connection reference may be had to some other experiments, in which vocal and manual processes were directly compared in somewhat more complicated relations than here involved. Briefly, certain cognitive and aesthetic judgments were timed, the responses being either with right and left hands on telegraph keys, or *yes* and *no* with the voice. Under these conditions, the excess of vocal over manual time becomes much smaller or negative. It is more natural to answer many questions by yes or no than by the manipulation of telegraph keys, and it takes less time. This would argue in favor of utilizing verbal responses even in the chronoscopic measurement of the more complex mental processes. But there are technical difficulties, chiefly in that voice-keys respond better to some voices than others. Electromyographic registration, as employed by Dodge, is perhaps the best solution, where the necessary equipment is available.

The idea of a motor factor in every mental process has been current in psychology for many years. Recently emphasis has been laid on vocal mechanisms as having a dominant, if not essential rôle in this motor factor. Under such circumstances, vocal terms would be the easiest in which mental processes could be expressed. This seems to be the case as the more complicated choice reaction processes are reached. With less complicated ones, however, of the level of distinguishing black and white, the advantage seems to lie rather with manual processes. In Subject IV., for example, the vocal times are twenty-five per cent. or more in excess of the manual ones. To postulate sub-vocal mechanisms in such a case would verge on schematism. This does not affect the concept that motor or 'behavior' factors are the essence of so-called mental activity. It does, perhaps, argue against supposing a vocal or sub-vocal mechanism, with all its close connections with mental activity, to be an essential channel, through which mental activity must take place.

CHARACTER TRAITS IN SCHOOL SUCCESS

BY A. T. POFFENBERGER AND FLORENCE L. CARPENTER

Psychological Laboratory, Columbia University

The relatively low coefficients of correlation (hovering around .50) between performance in intelligence tests and performance in school work are generally recognized to be due, not so much to defects in the tests as to the fact that school achievement involves other factors than those measured by means of the intelligence tests. A coefficient of correlation of .50 between record in an intelligence test and record in school work is a statistical expression for the fact that some children do better in school than their intelligence test records would indicate that they should, while others do not do as well as they should when judged on this basis, but that on the whole success in test and success in school go together. One form of discrepancy noted especially in our public school children has been explained as due to the tendency on the part of teachers to spur on the inferior child and to retard the superior child. But other causes must be sought for the fact that there are also some children who test low and do poorer school work than they should and some children who test high and do better work than they might be expected to do. Now the forces producing this condition are no doubt many and varied. Certainly health, home conditions, etc., would play a part. An important group of other factors is indicated by the term character traits.

It is our belief discussed in an earlier report of one of the writers [1] that it is quite feasible to construct an efficiency or competence test which shall measure at one operation the intelligence and the character traits vital for success. It is the purpose of this study to seek some of the character traits which are factors in school success and which should be included in such a competence test. Brown [2] has shown in a

[1] Poffenberger, A. T., 'Measures of Intelligence and Character,' *J. of Phil.*, 1922, *19*, 261–266.

[2] Brown, W. M., *Archives of Psychology*, No. 65, 1923.

recent study of Columbia Freshmen and other groups that a character trait which he has called 'Caution' may be readily measured by the Thorndike Intelligence Examination for High School Graduates and that this trait has an importance about one-third as great as intelligence in predicting success in school work.

The subjects chosen for this study were 97 school children in grades six, seven and eight. The makeup of the group is shown in the following table giving, for each of the three grades, their range in chronological age, range and average intelligence test score (measured by means of the National Intelligence Test, Scale *A*, Form 1), range and average I.Q.

Grade	Chron. Age	Test Score		I.Q.	
	Range	Range	Av.	Range	Av.
6	10 : 11 to 13 : 11	76 to 137	107	82 to 135	105.7
7	11 : 7 " 14 : 10	88 " 154	126	86 " 147	113.3
8	12 : 1 " 15 : 7	103 " 179	137	94 " 160	112.3

These 97 children were ranked for school success by each of three teachers, all equally familiar with each child's school performance. A final rank for each child was obtained by taking the average of the positions assigned by the three judges. In order to avoid the difficulty of making a rank order for so large a group the children in each grade were judged separately.

Each child could now be assigned a position in his group, on the basis of each of three sets of data; first, a position based on the average judgment of his teachers as to his school success; second, a position based on his performance in the intelligence test; and third, a position based on his I.Q. The relationship among these three rankings thus obtained may be seen from the following coefficients of correlation (Rank-Difference) between School Success and Test Score and between School Success and I.Q.

School Success and	Test Score	I. Q.
Grade 6	.70	.70
" 7	.74	.77
" 8	.69	.77

These coefficients are higher than are usually found. Their size may be accounted for, in part at least, by the range of intelligence score and I.Q. in each of the three grades and the fact that the number of cases is small.

The three rankings of each grade were carefully examined and two groups of children were selected for further study. First, those children were found who held a higher position in school performance than in either the test score or the I.Q. There were thirty three cases of this sort, and they form what will be called in this report the 'Success' group, meaning simply that each child in the group had a better school record than his intelligence test record or I.Q. would indicate. Second, those children were found who held a lower position in school performance than in either the test score or the I.Q. There were 34 children in this group which will hereafter be called the 'Failure' group. It must be remembered that there were no actual failures and that all the children were doing passable school work. The terms 'Success' and 'Failure' means simply better or worse than intelligence or I.Q. records warrant. The remaining cases, thirty in number, which had the same position in school performance as in either of the other two measures were discarded as being neither better nor worse than could be expected. The chronological ages of the 'Success' group and the 'Failure' group are as follows:

Grade	'Success' Group	'Failure' Group
6	12 : 9	12 : 7
7	13 : 6	13 : 3
8	14 : 4	14 : 10

The average I.Q. for the 'Success' group is 106.6 and for the 'Failure' group is 113.0. This means that the children of high I.Q. were more likely to do less than might be expected of them than were the children of lower I.Q. This bears out a current view previously mentioned that there is a tendency to let the more intelligent children work below their capacity as measured by the tests. In average chronological age there is no marked difference between the two groups.

The 'Success' and 'Failure' groups were given the Downey Will Temperament Test (as modified for group use by the

Bureau of Personnel Research of Carnegie Institute). Certain slight modifications were made in the instructions for the test in order that they might be more readily understood by the youngest children. No other changes in procedure or scoring were made. No total scores were computed, but instead scatter diagrams were compiled for the separate parts of the test in the manner described by Bills.[1] Each one of the twelve parts of the test was plotted separately against each other part, making a total of sixty six diagrams. These were inspected to find critical combinations of scores, or those that were characteristic of 'Success' or 'Failure' in school accomplishment. Thirty one of the pairs were found to give such combinations of scores, and these were used in our computations described below. The records of each individual in our two groups were examined, and whenever he had a score in any pair of traits which put him in the 'Success' area of a scatter diagram, he was given a score of + 1 for that combination of traits; whenever he had a score which put him in the 'Failure' area of a diagram he was given a score of − 1 for that combination of traits. If the scores obtained in this way by each individual child for the whole thirty one combinations are averaged separately for two groups, the following figures are obtained:

	'Success' Group	'Failure' Group
Av. of plus scores	9.4	3.5
Av. of minus scores	3.6	9.9

The method of making up the composite score is, indeed, crude, and it is quite possible that some more refined method, with proper weighting of different components of the Will Temperament test would reveal a much more pronounced difference between these two groups. It should be recalled in this connection that the coefficients of correlation between test score and school performance were rather high, so that the influence of the other factors involved, such as character traits must have been correspondingly slight. The above figures do suggest that the group of 'Success' children possess

[1] 'Method for the Selection of Comptometer Operators and Stenographers,' *J. of Applied Psychology*, 1921, *5*, 275–283.

certain combinations of scores in common, and that the 'Failure' group possess certain different combinations of scores in common.

	Speed of Movement		Motor Inhibition		Speed of Decision		Freedom from Load		Motor Impulsion		Flexibility		Perseverance		Care for Detail		Coördination of Impulses		Speed of Decision		Assurance		Self-Consciousness	
	H	L	H	L	H	L	H	L	H	L	H	L	H	L	H	L	H	L	H	L	H	L	H	L
Speed of Movement . H							+		+	−									+					
L							−	+												−				
Motor Inhibition . . . . . H					−		−								+	−		−						
L					+			+							+			+						
Speed of Decision . . . . . H			−	+					+	−			+	−							+			
L															+	−								
Freedom from Load . . . . . H	+	−	−							−			+	−				+		−		−		
L		+		+					+						+								+	
Motor Impulsion . . . . . H	+				+			+			+				+				−	+		+		
L	−				−		−					−				−							−		
Flexibility . . . . H									+							−						+	−	
L										−														
Perseverance . H					+		+															+		
L					−		−																	
Care for Detail . . . . . H				+		+		+	+											+	+			
L			−			−				−		−							−					
Coördination of Impulses H																			+	−				
L			−	+			+									−			−	+		+		+
Speed of Decision . . . . . H	+								−								+	−	−					
L		−					−		+						+	−	−	+						
Assurance . . . . H					+					+					+	+								
L							−		+	−	−							+						
Self-Consciousness . . H								+																
L																		+						

The accompanying chart summarizes in a schematic fashion the combinations of traits which appear to make for success and failure (as defined in this study). The twelve traits are indicated by short expressions carrying the meaning originally given to the parts of the test by Downey. A plus sign always means characteristic of 'Success,' and a minus sign means characteristic of 'Failure.' H means 'High Score' and L means 'Low Score.' For example, in the first row of the chart, we read that high 'Speed of Movement' and great 'Freedom from Load' make for success, while low 'Speed of Movement' and great 'Freedom from Load' make for failure. Furthermore, low 'Speed of Movement and lack of 'Freedom from Load' make for success.

An examination of this chart shows that possession of the following traits imply 'Success' in all the combinations in which they appear:

1. Great care for detail.
2. Lack of freedom from load.
3. Lack of motor inhibition.
4. High degree of assurance.
5. High degree of perseverance.

Trait (2) above seems to demand some explanation. The score in this test is the ratio of speeded writing to normal writing. Therefore, a low score (Lack of freedom from load) may indicate, not that the subject was unable to speed up under pressure, but rather that under normal circumstances his speed was near his maximum. This tendency to work at high speed without external pressure would obviously be a desirable trait. It is also likely that 'Lack of Motor Inhibition' shown by inability to retard the speed of writing to an extreme degree should be characteristic of the relatively successful students in a school.

The following traits were characteristic of failure in all the combinations in which they appear:

1. Lack of care for detail.
2. Lack of motor impulsion.
3. Motor inhibition.
4. Lack of flexibility.

The following traits were characteristic of success in all combinations except one as noted:

1. Motor impulsion except in combination with swift decision regarding self.
2. Flexibility except with low assurance.
3. High speed of movement except in combination with lack of motor impulsion.

The following traits were characteristic of failure in all the combinations except one as noted:

1. Low speed of movement except in combination with lack of freedom from load.
2. Lack of perseverance except in combination with high assurance.

The five pairs of traits whose possession is most indicative of success and the five whose possession is most indicative of failure are given below:

'Success'	*'Failure'*
Swift coördination and swift decision on self.	Slow coördination and swift decision on self.
Motor impulsion and flexibility.	Lack of both motor impulsion and flexibility.
High speed of movement and freedom from load.	Low speed of movement and freedom from load.
Care for detail and lack of motor inhibition.	Lack of care for detail and motor inhibition.
Flexibility and high assurance.	Flexibility and low assurance.

This brief study suggests two conclusions, each of which should be further tested upon large groups of subjects and with more refined technique:

1. Certain so-called character traits do contribute to success of failure in school work. These are, on the whole, what common sense would lead us to expect. For instance, speed and flexibility of reaction, assurance, perseverance and care for detail are traits indicative of success, while the failures tend to lack speed, and flexibility, to be careless and not

persevering. Other traits take on significance when one goes back of the mere name and analyzes the actual task performed. All such important traits should be discovered and an attempt made to determine their relative weight.

2. These traits are such as can be measured by relatively simple test methods, which might be incorporated into our present intelligence examinations with slight modifications. Care for detail, perseverance, freedom from load, and flexibility would seem to the writers to be most readily measured in this way.

CALCULATING THE STANDARD DEVIATION BY PLOTTING AND WITHOUT GROUPING THE RAW MEASURES

BY HERBERT A. TOOPS,

Institute of Educational Research,
Teachers College

There are several ways by which the standard deviation can be computed by using either the raw measures or positive 'steps' considered as raw measures, which will eliminate the undesirable negative deviations and separately calculated corrections for an assumed mean.

If the raw measures are small in magnitude so that the squares of the raw measures may be computed mentally (Barlow's Tables may be used) and if the number of cases, N, is small, then formula (1) is probably the easiest way of computing the standard deviation. When the raw measures are large in size, this formula becomes too laborious for general use.

Where X is a raw measure,

$$\sigma_X = \sqrt{\frac{N \cdot \Sigma X^2 - (\Sigma X)^2}{N^2}} \,. \qquad (1)$$

Formula (1) is to be used whenever the raw measures are small in size and few in number, whenever it is desirable that the raw measures be not grouped into classes, and *when the raw measures are decimals.* The use of this formula is shown in Table I., where the raw measures are so large as to make the squares of the raw measures unwieldy. Its advantage, where X is small, can be readily inferred from Table I. This method is not an approximation but will yield the standard deviation correctly to any desired number of decimal places, an advantage not possible by formulæ which employ

d_x corrections, or where the raw measures are grouped into classes.

Table I

Computing the Standard Deviation by Formula 1

(These raw measures are here given at length as they are later used in checking the plotting method.) $N = 50$ cases.

X	X^2	X	X^2	X	X^2	X	X^2	X	X^2
6	36	43	1849	56	3136	67	4489	84	7056
12	144	44	1936	58	3364	70	4900	86	7396
15	225	44	1936	58	3364	72	5184	88	7744
18	324	45	2025	58	3364	74	5476	88	7744
21	441	47	2209	58	3364	74	5476	88	7744
25	625	49	2401	62	3844	76	5776	91	8281
28	784	50	2500	65	4225	76	5776	93	8649
35	1225	51	2601	65	4225	79	6241	95	9025
41	1681	51	2601	67	4489	82	6724	97	9409
43	1849	56	3136	67	4489	82	6724	98	9604

$\Sigma X = 2998 \quad 207{,}810 = \Sigma X^2$

$$\sigma_X = \sqrt{\frac{50(207{,}810) - (2998)^2}{(50)^2}} = \sqrt{\frac{1{,}402{,}496}{2500}} = 23.685405.$$

In an actual situation, with data not intended for illustration as is the above, the raw measures do not come arranged in order, nor is such order necessary. All of the work may be done on the adding machine.

The formula may be derived from the usual Pearson formula by making the substitution

$$\left(X - \frac{\Sigma X}{N}\right) = x.$$

When the number of cases is very large, several hundred cases say, the method becomes quite too laborious for practical use even if the raw measures are small in size. The following plotting method is efficient when N is very large, and when the raw measures can be plotted within the limits of the table used.[1]

We may plot the raw measures in a chart such as Fig. 1, where the abscissæ of the chart represent the tens digits and the ordinates the units digits of the several raw measures

[1] The author's latest revision of correlation chart allows one to compute σ, obtaining full value of the variability of raw measures between 0 and 179 inclusive, by letting a take the values 0 to 17 inclusive.

Tens (a).

$Fr_D(a-b)^2$	0	64	98	180	75	96	81	16	15	625 C
$(a-b)^2$	81	64	49	36	25	16	9	4	1	
Fr_D	0	1	2	5	3	6	9	4	15	5 — 50 N
	I	H	G	F	E	D	C	B	A	L.D.
	0	1	2	3	4	5	6	7	8	9

Units (b).

											Fr_b	b	$Fr_b \cdot b$	$b(Fr_b \cdot b)$	b^2
9					I			I			i	2	9	18	162	81
8		I	I			IIII			III	I	h	10	8	80	640	64
7					I		III			I	g	5	7	35	245	49
6	I					II		II	I		f	6	6	36	216	36
5		I	I	I	I		II			I	e	7	5	35	175	25
4					II			II	I		d	5	4	20	80	16
3					II					I	c	3	3	9	27	9
2		I					I	I	II		b	5	2	10	20	4
1			I		I	II				I	a	5	1	5	5	1
0						I		I			*	2	0	00	000	0

	*	a	b	c	d	e	f	g	h	i	
Fr_a	1	3	3	1	8	9	6	7	7	5	50 N — 248 E — 1570 B
a	0	1	2	3	4	5	6	7	8	9	
$Fr_a \cdot a$	00	3	6	3	32	45	36	49	56	45	275 D
$a(Fr_a \cdot a)$	000	3	12	9	128	225	216	343	448	403	1789 A
	0	1	4	9	16	25	36	49	64	81	

$$\sigma_x = \sqrt{\frac{\binom{N}{50}\left[110\binom{A}{1789} + 11\binom{B}{1570} - 10\binom{C}{625}\right] - \left[10\binom{D}{-275} + \binom{E}{248}\right]^2}{\binom{50}{N}^2}}$$

$$= \sqrt{\frac{1{,}402{,}496}{2500}} = 23.685405$$

4-15-28.

—Herbert A. Toops.—

respectively. Thus, 97 is plotted as a tally mark in the 9-column and 7-row; 6 is plotted in the 0-column and 6-row; 25 in the 2-column and 5-row, and so on.

Let a represent the tens digit, and

b represent the units digit of a given raw measure, X.

Then

$$X = (10a + b), \tag{2}$$

whence

$$\Sigma X = 10\Sigma a + \Sigma b \tag{3}$$

and

$$\Sigma X^2 = 100\Sigma a^2 + 20\Sigma ab + \Sigma b^2. \tag{4}$$

We may now eliminate the undesirable Σab term as follows:

$$\Sigma(a - b)^2 = \Sigma a^2 - 2\Sigma ab + \Sigma b^2. \tag{5}$$

Solving for Σab,

$$\Sigma ab = \frac{\Sigma a^2 + \Sigma b^2 - \Sigma(a - b)^2}{2}. \tag{6}$$

Substituting equation (6) in equation (4), and collecting terms,

$$\Sigma X^2 = 110\Sigma a^2 + 11\Sigma b^2 - 10\Sigma(a - b)^2. \tag{7}$$

We may now substitute equations (7) and (3) respectively in equation (1), whereupon,

$$\sigma_X = \sqrt{\frac{N[110\Sigma a^2 + 11\Sigma b^2 - 10\Sigma(a-b)^2] - [10\Sigma a + \Sigma b]^2}{N^2}}. \tag{8}$$

All of the quantities of the above formula may be obtained from the printed correlation sheets used by the author.[2] The quantity, $\Sigma(a - b)^2$ may be obtained by counting diagonal frequencies. The raw measures are to be plotted as if one were computing the correlation between the tens digit (X) and units digit (Y) of the raw measures. All sums are to be found as in solving this correlation, merely by making the substitutions for the X and Y notation of the correlation chart.

[2] 'Eliminating the Pitfalls in Solving Correlation: A Printed Correlation Form,' Jour. of Exp. Psychology, Vol. 4, No. 6, 1921, pp. 434–446.

$$a = X,$$
$$b = Y,$$
$$\Sigma a = \Sigma X = D,$$
$$\Sigma b = \Sigma Y = E,$$
$$\Sigma a^2 = \Sigma X^2 = A,$$
$$\Sigma b^2 = \Sigma Y^2 = B,$$
$$\Sigma(a - b)^2 = \Sigma(X - Y)^2 = C.$$

This formula then becomes,

$$\sigma_x = \sqrt{\frac{N[110A + 11B - 10C] - [10D + E]^2}{N^2}}. \qquad (9)$$

In the example, Fig. 1, where the raw measures of Table I. are plotted, this becomes,

$$\sigma_x = \sqrt{\frac{(50)[110(1789) + 11(1570) - 10(625)] - [10(275) + (248)]^2}{(50)^2}}$$

$$= \sqrt{\frac{1,402,496}{2500}} = 23.685405.$$

The answer can obviously be carried to any number of decimal places as the decimals do not enter until the last operation.

Comparison of the New Method with the Customary Grouping Method

When the raw measures are grouped in steps of 0–9, 10–19, 20–29, 30–39, etc., the standard deviation can be found by the formula,

$$\sigma_x = I_x \sqrt{\frac{N\Sigma a^2 - (\Sigma a)^2}{N^2}}, \qquad (10)$$

where I_x is the class interval, here 10. By substituting the values of Fig. 1, the standard deviation becomes

$$\sigma_x = 10\sqrt{\frac{13,825}{2500}} = 23.515952.$$

Formula (10) makes use of positive steps of the classified or grouped measures. It likewise contains no decimals until

the final operation, and so yields the correct σ for the *grouped measures* considered as positive raw measures; such grouping always introduces the "grouping error," represented in this case by the difference between the standard deviation of formula (8) and that of formula (10) or 7-10 of 1 per cent of error in this case. The true σ is given by both formulæ (1) and (8). By making I_x smaller, and thus using more classes, the grouping error is reduced.